Diagnosis of Total Hip Prosthesis Condition Using a Combined Approach of Acoustic Emission Monitoring and Patient Gait Analysis

Anthony J. FitzPatrick

A thesis submitted in partial fulfilment

of the requirements for the Degree

of

Doctor of Philosophy

in

Mechanical Engineering

at the

University of Canterbury,

Christchurch, New Zealand

July 2019

Abstract

An increasing demand for total hip replacement (THR) surgeries, due to an ageing population and an increasing number of younger patients undergoing THR surgery, has resulted in a proportional increase in the number of patients requiring revision surgery to remedy complications or failures of their THR implant components. Revision surgery of THR implants is expensive and time consuming and places an increasing burden on health funding and surgical resources. Consequently, a large emphasis has been placed on developing more accurate diagnostic methods which allow early diagnosis of THR implant wear and complications. This thesis investigates the potential of acoustic emission (AE) monitoring as a method for the early diagnosis of THR implant wear and complications.

A detailed data collection protocol is presented for the use of an AE monitoring device on THR patients in the clinical environment (*in-vivo*) and on implants retrieved from THR patients in the laboratory environment (*in-vitro*). The AE monitoring device samples AEs at 100 kHz using ultrasonic receivers placed on the skin surface over the hip joint of THR patients. During *in-vivo* monitoring, THR study participants perform basic motions such as walking, sitting to standing, and stair ascents and descents. *In-vitro* data collection utilises the same AE sensing equipment to record AEs from robotic and manual manipulations of THR implants retrieved during the revision surgery of study participants.

A detailed analysis was performed on the time and frequency domain characteristics of the *in-vivo* AEs collected from 117 THR patients and 23 control study participants. *In-vitro* data was collected and analysed from the retrieved implants of five THR patients with ceramic implant bearing components. Dominant frequencies of *in-vivo* AE signals were found to be in the 1-4 kHz range. *In-vitro* AEs from implant main bearing surface manipulations yielded dominant

frequency content in the same 1-4 kHz range. However, *in-vitro* AEs from relative motion at the trunnion Morse taper of the femoral implant components yielded squeaking AEs with dominant frequency content in the 18-22 kHz range.

Further work developed and implemented an event detection algorithm to automate the detection of regions of interest (AE events) from the large dataset of AE signals. The AE events from patients who progressed onto revision surgery were analysed and grouped using the clinically identified reasons for revision of excessive noise, loosening, and excessive wear. A statistical analysis of AE event parameters was performed on the grouped data and the analysis indicated that single parameters alone were not an effective way to categorise participant AE data. A multidimensional analysis of the event parameters was performed, which took the form of two-parameter scatter plot comparisons and a principal component analysis. The multidimensional analyses showed that AEs were generally similar across all groups. However, some definite differences in the distributions of AE events from the different groups was apparent, particularly those within the loosening group of THR participants. The overall results from the AE analysis showed promise in the diagnostic potential of the AE monitoring technique. However, the results fell well short of being able to lead to a reliable failure diagnosis in their current state.

Finally, a system to concurrently collect combined AE and gait data from THR patients was developed. The combined system provides the ability to continuously link the AE signals with the specific motion of the patient, and therefore to the motion of the implant components. The system itself and detailed data collection protocols were developed and combined AE and gait data was collected and analysed from three THR study participants. The data from the combined system has shown promising intra- and inter-participant consistency with substantial AE activity observed to occur repeatedly during a subset of the stance phase of the gait cycle.

Audible squeaking was also observed to occur consistently during the terminal stance phase for one of the participants, repeatedly occurring within the 30-40% range of the gait cycle. The initial testing from the combined AE and gait analysis has indicated relationships likely exist between AEs and particular implant motions and loadings.

While further research is needed before the AE monitoring technique could be used as a clinical diagnostic method that might augment existing methods, this thesis presents some key steps towards this final overarching goal. Critical assessment of the current methods is made, and recommendations for key future research are presented.

Deputy Vice-Chancellor's Office
Postgraduate Office

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Co-Authorship Form

This form is to accompany the submission of any thesis that contains research reported in co-authored work that has been published, accepted for publication, or submitted for publication. A copy of this form should be included for each co-authored work that is included in the thesis. Completed forms should be included at the front (after the thesis abstract) of each copy of the thesis submitted for examination and library deposit.

Please indicate the chapter/section/pages of this thesis that are extracted from co-authored work and provide details of the publication or submission from the extract comes:

Chapters 3 & 4

FitzPatrick A. J., Rodgers, G. W., Hooper, G. J., Woodfield, T. B. F.

Development and validation of an acoustic emission device to measure wear in total hip replacements in-vitro and in-vivo

2017, Biomedical Signal Processing and Control, 33, 281-288. DOI: 10.1016/j.bspc.2016.12.011

Section 4.3.3

FitzPatrick A. J., Rodgers, G. W., Hooper, G. J., Woodfield, T. B. F.

Cepstrum Analysis for Determining Fundamental Frequency of Total Hip Replacement Acoustic Emissions

2017, International Federation of Automation and Control (IFAC) World Congress, Toulouse, France. DOI: 10.1016/j.ifacol.2017.08.1642

Please detail the nature and extent (%) of contribution by the candidate:

The contribution of the candidate to the publications amounts to ~85% based on the overall time and effort dedicated to the final outcome. The candidate contributed to data collection and formalised the data collection protocols. The candidate developed and implemented all data analysis and consulted with supervisors on the presentation methods. The candidate wrote the manuscripts, with supervisors providing feedback on the draft versions before submission. The candidate undertook the revisions to the manuscript based upon the reviewer comments, with input from supervisors.

Certification by Co-authors:

If there is more than one co-author then a single co-author can sign on behalf of all

The undersigned certifys that:

- The above statement correctly reflects the nature and extent of the PhD candidate's contribution to this co-authored work
- In cases where the candidate was the lead author of the co-authored work he or she wrote the text

Name: *Geoff Rodgers* Signature: *Geoffrey W Rodgers* Date: *29 / 7 / 2019*

Deputy Vice-Chancellor's Office
Postgraduate Office

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Co-Authorship Form

This form is to accompany the submission of any thesis that contains research reported in co-authored work that has been published, accepted for publication, or submitted for publication. A copy of this form should be included for each co-authored work that is included in the thesis. Completed forms should be included at the front (after the thesis abstract) of each copy of the thesis submitted for examination and library deposit.

Please indicate the chapter/section/pages of this thesis that are extracted from co-authored work and provide details of the publication or submission from the extract comes:

Chapter 4

Roffe, L., FitzPatrick A. J., Rodgers, G. W., Woodfield, T. B. F., Hooper, G. J.

Squeaking in ceramic-on-ceramic hips: No evidence of contribution from the trunnion morse taper.

2017, J Orthop Res, 35(8), 1793-1798. DOI: 10.1002/jor.23458

Please detail the nature and extent (%) of contribution by the candidate:

The contribution of the candidate to the publication amounts to ~70% based on the overall time and effort dedicated to the final outcome. The candidate contributed to data collection. The candidate developed and implemented all technical data analysis and consulted with supervisors. Note the candidate is considered equal first author and contributed ~50% to overall manuscript writing time. The candidate contributed 100% to writing of manuscript methodologies and results and contributed to writing of introductions and discussions. The named first author (Roffe, L.) was an orthopaedic registrar who conducted collation of patient clinical data and contributed to the text about clinical research significance and the associated discussion. It is important to note that the orthopaedic registrar did not have an engineering/technical background and did not contribute to the engineering analyses presented within this manuscript.

Certification by Co-authors:

If there is more than one co-author then a single co-author can sign on behalf of all

The undersigned certifys that:

- The above statement correctly reflects the nature and extent of the PhD candidate's contribution to this co-authored work
- In cases where the candidate was the lead author of the co-authored work he or she wrote the text

Name: *Geoff Rodgers* Signature: *Geoffrey W Rodgers* Date: *29 / 7 / 2019*

Acknowledgements

I would like to express my sincere gratitude to all who played a part in the successful completion of my Ph.D. degree. While I will only name a few here who helped in concrete ways, I am eternally thankful for all those who offered friendship, encouragement, and prayers along the way.

I thank God for this accomplishment, without His quiet influence in my life I can see the path would have been much more difficult.

I would like to express my sincerest thanks to my main supervisor, Geoff Rodgers, who gave me tireless support throughout the entire process, especially by means of invaluable advice, practical assistance with all data collection, and through obtaining the Marsden Fund and MedTech CoRE grants which funded this research.

I must also thank my co-supervisors, Gary Hooper, Justin Fernandez, and Tim Woodfield, for their role in my success. I especially wish to thank Gary for his expert guidance on clinical matters, and Justin for passing on his invaluable experience with gait tracking and analysis.

I want to thank Lyn from the COBRA unit at Burwood hospital for liaising with all the study participants and making the clinical testing sessions possible. I also thank Gerry Kirk for his knowledge and support in the software development process. I thank Lloyd and Sam for searching the patient records and collating the clinical data.

I am also immensely thankful to all my friends and colleagues who made the process extremely enjoyable. I must give a special thanks to Arun for his assistance with equipment testing, data collection, and joyful conversations about all things life. I must also especially thank my good friend, Carmela, for making the data collection process a success by agreeing to be the test subject for the trial run.

Finally, I wish to extend a special thanks to my parents and to my best friend, Jemima. I am forever grateful for their endless love, support, and encouragement of my endeavours and I would not be who I am today without their loving influence.

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Total Hip Replacement Incidence

The leading cause for primary total hip replacement (THR) surgeries in many countries is reported to be osteoarthritis, also known as degenerative joint disease (Australian Orthopaedic Association National Joint Replacement Registry, 2015; Canadian Institute for Health Information [CIHI], 2015; Garellick et al., 2015; Hooper, 2013; National Services Scotland, 2016; New Zealand Orthopaedic Association [NZOA], 2015; Pivec et al., 2012). Osteoarthritis causes pain and loss of normal function in joints and is the most common joint disease worldwide with an estimated 10% of men and 18% of women over the age of 60 being affected (Glyn-Jones et al., 2015). When osteoarthritis is found to be at an advanced stage often the only effective treatment available is replacement of the joint with an artificial implant. With demographic ageing of the population, the incidence of osteoarthritis and consequent THR surgery is becoming increasingly more common in New Zealand and around the world.

In 2016 there were 8785 primary THR surgeries in New Zealand which was 5% more than in 2015 and more than double the amount of surgeries in 1999 when the registry began (NZOA, 2017). Projections for the next 20 years suggest a threefold increase in the number of THR surgeries in the USA for patients over 65 years of age and data from several other countries show similar trends (Hooper, 2013; Hooper et al., 2014; Kurtz, S. et al., 2007). Additionally, trends from around the world show an increasing proportion of THR patients are below the age of 65 and this proportion is approaching or has exceeded 50% (Kurtz, S. M. et al., 2009; NZOA, 2015). As a result of an increasing number of younger patients undergoing THR surgery, joint replacement implants are being required to remain in service longer and withstand a more active patient lifestyle.

An initial THR surgery to implant an artificial joint, known as primary surgery, has a high success rate returning patients to pain-free activities. In New Zealand THR implants have a 15 year survival rate of 87.3% and data from the UK and USA shows a greater than 80% survival rate after 25 years (NZOA, 2015; Pivec et al., 2012). However, despite a high survival rate, the increase in the number of primary THR surgeries means that the number of patients who develop implant complications or failures is also increasing substantially. As a consequence, the demand for revision surgery (a new operation where a component is exchanged or manipulated in a previously replaced hip joint) is continuing to increase proportionately. Revision surgery is expensive and time consuming and places an increasing burden on health funding and surgical resources: The cost of a THR revision can be up to four times that of the primary procedure (Wyatt et al., 2014). With the knowledge of an ever increasing number of revision surgeries, an emphasis is being placed on improving reliability and survivorship of joint replacement prostheses (Kurtz, S. M. et al., 2009).

1.2 Research Motivation

One step towards the goal of improving the reliability and survivorship of joint replacement implants is by placing emphasis on developing more accurate diagnostic methods which allow early diagnosis of THR prosthesis wear and complications (Hayter et al., 2012; Rodgers et al., 2014). If the condition of THR prostheses can be easily and reliably monitored more frequently, it is hoped that many major revision surgeries can be avoided or reduced in complexity through the early detection of complications and wear. As a consequence, the overall cost and resource burden for the health system can be reduced and implant manufacturers may gain better insight into implant factors that initiate failure.

Whilst many implant complications will require surgery to remedy the situation, early detection can often lead to a less invasive surgery where components interfacing with bone are left in place and only bearing surface component exchanges occur. A less invasive surgery is advantageous for patient recovery and improves the chance of long term success of the implant. If failing implants are not detected, wear debris from bearing surface degradation causes osteolysis (pathological destruction of bone around the implant from the body's reaction to foreign wear particles). The degree of osteolysis increases with time, reducing the bone stock available for future implant fixation, significantly complicating the revision surgery and reducing the chance of success (Dattani, 2007). Early diagnosis can also provide further valuable information to surgeons and prosthesis manufacturers and allow them to improve surgical procedures and prosthesis designs in an attempt to reduce the need for revision surgery.

There are many techniques that can be used in the diagnosis of THR complications, all with various advantages and disadvantages. There is much literature to suggest that no current single technique provides a definitive clinical diagnosis for all of the different THR complications. Therefore, emphasis is currently placed on using outcomes from sequential changes in radiographs and multiple imaging techniques coupled with detailed patient histories and physical examinations (Burge, 2015; Clohisy et al., 2009; Miller, 2012). It is evident that X-ray imaging is the most common diagnostic technique currently used in THRs (outside physical examination) due to its relative low cost, ease to produce, and its ability to provide good preliminary insight into prosthesis condition. If further diagnostic information is required CT and/or MRI are usually employed and, if necessary, other techniques can also be used such as fluoroscopy (a technique which displays real-time X-rays to allow viewing of moving structures radiographically).

However, the major concern with the use of X-ray imaging, CT scanning, and fluoroscopy is the potential for tissue damage by ionizing radiation. Hence, research into new diagnostic techniques that are non-invasive, inexpensive, simple to conduct, and free from ionizing radiation has great importance. One such technique of particular interest, and the subject of this thesis, is the detection and analysis of acoustic emissions from hip replacement implants. This technique is not currently employed clinically in orthopaedics and there is limited literature on the subject.

The remainder of this chapter will briefly introduce the technique of acoustic emission monitoring and then go on to explain the previous work surrounding acoustic emission monitoring of THR implants that has been conducted at the University of Canterbury and the University of Otago Christchurch. This previous work forms the basis for the work that was conducted for the present thesis. The final sections will outline the research objectives and provide a chapter overview of this thesis.

1.3 Acoustic Emission Monitoring

An acoustic emission (AE) is considered a mechanical elastic wave that propagates through a medium, or multiple mediums, as a result of dynamic local rearrangements at an atomic scale caused by strain in the material from which the wave originated (Baranov et al., 2007; Eitzen & Wadley, 1984; Kohn, 1995; Scruby, 1987). The atomic rearrangements are a result of the conversion of elastic or kinetic energy, from an applied stimulus, to thermal energy through the local deformation, or atomic vibration and rearrangement, of a material in order to dissipate the energy (Akay, 2002). It then stands to reason that, because an AE concerns vibrations at an atomic level, AEs will be present in all physical or dynamic interactions of a substance and knowledge of the presence of AEs is only limited by the sensitivity of equipment used to detect

them (Baranov et al., 2007). AEs are typically detected using transducers attached to the surface of the material in question. The transducers detect the surface motion, which is a result of the mechanical vibration of the AE, and convert it to a voltage-time waveform allowing the characteristics of the AE to be investigated (Eitzen & Wadley, 1984).

AE monitoring has been considered an ideal damage monitoring technique due to its sensitivity to deformation processes and its ability to be applied continuously (Kohn, 1995). It can be used to non-invasively and non-destructively monitor both surface and subsurface deformation, fatigue, and fracture of various materials, including soft tissues and metal or ceramic components. The AE technique can also provide insight into the materials' failure mechanisms and over the past half century the technique has advanced considerably. Literature indicates the technique has produced some promising results in orthopaedics by attempting to diagnose various prosthesis complications such as loosening, osteolysis, and infection with varying degrees of success. However, the literature on the subject is both fragmented and limited: most of the studies generally do not have the aim of using the AE data to make a clinical diagnosis and the studies that do have uncorroborated results. A detailed review of the existing literature specifically surrounding the use of AE monitoring in orthopaedics, including findings and limitations, is presented in Section 2.3.2.

1.4 Previous Work

1.4.1 Overview

Previous work at the University of Canterbury has developed an AE prototype diagnostic tool for joint replacement assessment and clinical trials have been conducted, in conjunction with the University of Otago Christchurch School of Medicine and the Canterbury Orthopaedic and Bone Research Association (COBRA), to gather AE data from THR patients in the clinical (*in-*

vivo) environment (Rodgers et al., 2014). In addition to this, data has also been collected from laboratory (*in-vitro*) testing using THR implants retrieved from patients who have undergone revision surgery and were also previously monitored for AEs *in-vivo*. To our knowledge, the study is the first to be utilising both *in-vivo* and *in-vitro* AE data from the same THR implant components. The study therefore has the advantage of attempting to replicate implant specific *in-vivo* AE responses in the laboratory, as well as physically examining the condition of the implant components. By doing this, a more clinically relevant database of THR AE characteristics can be developed and a greater understanding of inter-patient variability can be gained. These factors should ultimately lead to an instrument with greater diagnostic potential than is currently available.

1.4.2 Acoustic Emission Device & Signal Processing

A simple sensing device had been developed which recorded AE data signals using National Instruments (NI) data acquisition hardware (CompacDAQ and NI-9222 analogue module) and LabVIEW software. The device utilised four passive ultrasonic receivers (Prowave 328ER250 sensors) each with a resonant frequency of 32.8 kHz. Each of the four data signals were recorded simultaneously at 100 kHz, giving AE signals with a maximum observable (Nyquist) frequency of 50 kHz. The LabVIEW software program output the recorded AE signal data as text files which were saved for post-processing.

Software code had been developed in MATLAB to import and organise the recorded AE signals to allow further analysis to be performed. Additionally, the MATLAB program had been designed to filter any unwanted signal noise from the data using a Daubechies wavelet filter. The software programming up to the point of producing organised and filtered AE data signals was considered both necessary and reliable and therefore will be used within the current

research and will not be modified. Some programming work had also been carried out on AE categorisation ('event detection') and spectral analysis but had yielded limited results. It had however allowed for some interesting findings relating to squeaking of THR implants and had demonstrated the generally varied nature of biological data. It was evident that more work was needed to effectively categorise the AE data. Long term, it would be desirable to develop real-time processing algorithms to process data and provide immediate feedback to an orthopaedic surgeon. However, this thesis will focus on developing reliable post-processing algorithms and hence real-time processing is considered outside of the scope.

1.4.3 Acoustic Emission Data Collection

A protocol for *in-vivo* AE data collection using the AE sensing device had been developed where the four sensors are attached against the skin surface of the participant from the iliac crest to the upper-femur. Ethical approval for *in-vivo* AE data collection was granted from the New Zealand Upper South A regional ethics committee under approval number URA/10/11/075. The participant performs a range of prescribed motions and any AEs that are produced which reach the skin surface are recorded. The study had previously collected AE data for 80 THR study participants, 39 of which have undergone subsequent revision surgery with 19 of those making their implant components available for *in-vitro* examination. AE data for 22 natural hips with no known previous joint problems or trauma had also been collected for use as control data.

An *in-vitro* test rig had been developed whereby a robotic arm is used to manipulate retrieved prosthetic joints in a reliable and repeatable manner. Prior to the inclusion of the robotic arm the *in-vitro* testing was carried out by hand and thus the robotic arm was incorporated in order to replicate the *in-vivo* movements in a more realistic way. The *in-vitro* testing set-up had not

been designed to replicate implant fixation to bone structures nor did it include any representation of soft tissues surrounding the implant. The main purpose of the *in-vitro* testing had been to investigate the interactions at the main bearing interface and the modular connections of the implant components. The AE sensing device had been adapted to record AEs generated by the implant components during *in-vitro* testing and to monitor unwanted external AE sources such as the actuating robot arm.

1.4.4 Surface Analysis

A number of the retrieved implant components, in particular some ceramic femoral heads, showed considerably worn surfaces. To investigate this wear more closely a surface analysis had been conducted at a microscopic level using scanning electron microscopy (SEM). Four ceramic implant components, which had been AE monitored both *in-vivo* and *in-vitro*, had subsequently been carbon coated and examined with a SEM. The analysis found a previously unobserved failure mechanism of inter-granular failure of the ceramic material which had been identified across all four components.

1.4.5 Modal Analysis

THR implant vibration modes had also been investigated using finite element analysis (FEA) software. A complete set of a THR implant had been laser scanned to yield solid models for use with ANSYS FEA software. A simple modal analysis had been carried out for each component to investigate the first three vibration modes and their frequencies. It became evident from the results that boundary conditions were an important factor in determining the modal properties and hence it was not expected that the FEA analysis frequencies would align accurately with the frequencies of the AEs. However, the results did show that the modal frequencies of the femoral head were considerably higher than those of the femoral stem,

acetabular cup, and acetabular liner. The modal frequencies of the stem, shell, and liner were more closely aligned with the main frequency content of *in-vivo* AEs as well as *in-vitro* AEs from the main bearing interface.

1.5 Research Objectives

The primary objective of this project was to use previously collected AE data from THR implants, along with any new data collected using the existing prototype diagnostic tool, to identify AE characteristics that are indicative of specific implant complications such as loosening or excessive wear. By identifying specific AE characteristics and correlating them with clinical reasons for revision, a significant step towards implant condition diagnosis can be made. In conjunction with this, the project was to continue building the existing AE database through ongoing clinical and laboratory data collection. A more representative and robust dataset was to be established to aid in clinical diagnosis through the collection of additional data.

A further objective was to design and implement a simple, quantitative, *in-vivo* gait analysis technique, using IMUs and 3D motion capture, to couple temporal planar implant joint angle data with corresponding AEs and, in turn, implant surface analysis. The existing *in-vivo* data collection protocol was limited in that it allowed only broad representation of the joint's motion over time during AE monitoring. For example, in a particular AE recording where a patient was walking, the general motion of the implant components was known but no temporal information of the implant orientations was recorded. Without any temporal implant orientation information the AE data cannot be associated with any specific implant positions. The additional information of temporal joint angles could offer further insight into the mechanisms and occurrence of AEs in joint replacements.

The hypothesis of this research is that THR implant complications have unique and identifiable AE characteristics that could be used to assist orthopaedic surgeons in making more informed and prompt implant complication diagnoses. From this hypothesis, and analysis of previous work, the following research questions were formulated:

1) Can both intra and inter-patient AEs from THR implants be reliably categorised using time and frequency domain characteristics?
2) If so, do these categories correspond to clinical reasons for revision such as loosening, noise, infection, or excessive wear?
3) Can any correlations be seen between implant joint angles and audible AEs?
4) If so, do these correlations allow more definitive conclusions to be drawn with respect to research questions 1) and 2)?

1.6 Chapter Overview

Chapter 2 presents the relevant background information and literature on three key topics: anatomical terminology and hip replacement implant components, AEs in medicine, and technical aspects of the AE technique and AE signals.

Chapter 3 presents the details of the AE monitoring hardware and data collection protocols for the *in-vivo* and *in-vitro* AE testing.

Chapter 4 presents the initial results of AE testing and familiarises the reader with the basic time and frequency domain appearances of the recorded AE signals from both the *in-vivo* and *in-vitro* testing.

Chapter 5 presents the AE event detection algorithm and subsequent evaluation. The chapter then goes on to show the results from the comparisons of the detected AE events of the control, noise revision, loosening revision, and excessive wear revision participant groups.

Chapter 6 presents the methodology and initial results for the developed technique of combined AE monitoring and gait analysis.

Chapter 7 presents the overall conclusions of the thesis and identifies future works related to the research.

1.7 Summary

This chapter presented the motivation for the research and introduced the concept of AE monitoring as a possible diagnostic tool for THR implants. An increasing number of primary THR surgeries are being performed due to ageing populations and an increase of younger patients undergoing THR surgery. This has resulted in a proportional increase in the number of THR patients experiencing implant complications and subsequent revision THR surgeries. Therefore, much emphasis is being placed on research into more accurate diagnostic methods that allow early indications of THR prosthesis wear and complications. Early diagnosis of implant complications can result in the avoidance of many major revision surgeries or reductions in their complexity. Consequently, the overall cost and resource burden for the health system can be reduced and recovery times for the patient could be improved. Implant manufacturers may also gain better insight into implant factors that initiate failure. Furthermore, diagnostic methods that are non-invasive, inexpensive, simple to conduct, and free from ionizing radiation are most desirable, such as the detection and analysis of AEs.

Literature on AEs in THR patients is limited but outcomes from the few published studies indicate the AE technique has produced some promising results in orthopaedics by attempting to diagnose various prosthesis complications such as loosening, osteolysis, and infection with varying degrees of success. However, most of the studies in the literature generally do not have the aim of using the AE data to make a clinical diagnosis and the studies that do have this aim have uncorroborated results. Therefore, the main objective of this thesis is to investigate AEs from a large number of THR patients and determine if the AEs have unique and identifiable AE characteristics that could be used to assist orthopaedic surgeons in making more informed and prompt implant complication diagnoses.

This chapter also presented the previous work that had been carried out to develop the current AE monitoring device and conduct initial data collection. A simple sensing device had been developed which recorded AE data signals at 100 kHz using ultrasonic receivers attached to the skin surface of THR study participants. Following ethical approval, AE data had been collected from 80 THR study participants and 22 control hips. In-vitro testing, on implants previously tested in-vivo, had also been conducted on implant components that were retrieved during revision surgery and subsequently made available for testing through informed consent by the associated study participants.

Chapter 2 Background & Literature Review

2.1 Introduction

In this chapter, relevant background information is presented on three key topics: anatomical terminology and the THR implant; literature of AEs in medicine; and technical aspects of the AE technique and AE signals. The information presented on the aforementioned topics lays the foundation for understanding the AE analysis that follows in the remainder of the thesis. The first sections of this chapter cover some relevant anatomical terminology that will be used throughout this thesis. Furthermore, an introduction to the hip joint and the THR implant equivalent is presented. Additionally, an explanation of some common THR implant failure modes and their associated prevalence is presented. The middle sections of this chapter present the literature concerning AEs in medicine and more specifically, the AE technique in the field of orthopaedics. The final sections of this chapter cover some technical background of the AE monitoring technique in general. In addition, some background concerning AE propagation and signal parameters is presented.

2.2 Medical Background Information

2.2.1 Anatomical Terminology & Hip Movements

To more effectively communicate particular aspects of this thesis it is useful to introduce some medical terminology that is used to describe the position and movements of parts of the human body. Figure 2.1a shows a diagram depicting three planes that are commonly used as references to describe relative positions or directions. In the context of this thesis, the sagittal plane is an important plane to note; the sagittal plane is a vertical plane that divides the body (or organ)

into left and right sides. The sagittal plane that divides the body into equal left and right sides is called the midsagittal plane. Figure 2.1b shows a diagram depicting the directional reference terminology used when referring to relative positions of body parts. The definitions of the directional terms that are useful to know are summarised in Table 2.1.

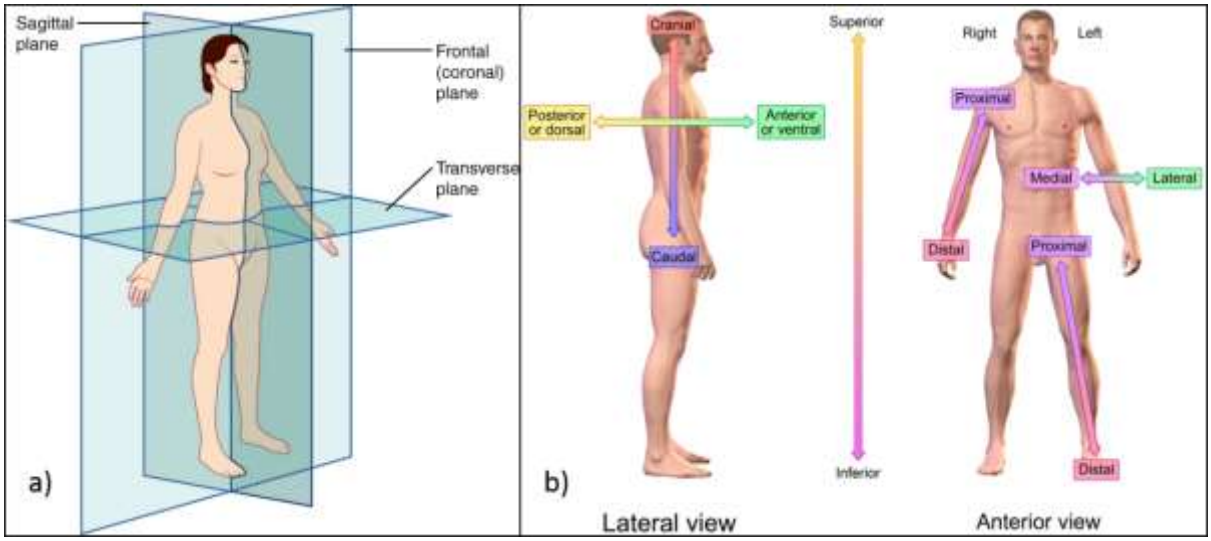

Figure 2.1: *a) Planes of the Body by OpenStax (2017, p. 6) licensed under CC BY 4.0,*
b) Directional References by Blausen.com staff (2014) licensed under CC BY 3.0.

Table 2.1: Definitions of body directional terms.

Directional Term	Definition
Anterior	Toward or at the front of the body
Posterior	Toward or at the back of the body
Superior	Toward the head, or the upper part
Inferior	Away from the head, or the lower part
Medial	Toward or at the midline (the midline is a line on the midsagittal plane)
Lateral	Away from the midline
Proximal	Toward or at the end of a limb where it attaches to the body
Distal	Away from or the end of a limb where it attaches to the body

Specific terminology is used for common movements that occur at particular joints of the body to indicate the form and direction of the motion. Figure 2.2 depicts the main movements of the hip joint and the associated terminology. A description of each of the possible hip movements is summarised in Table 2.2.

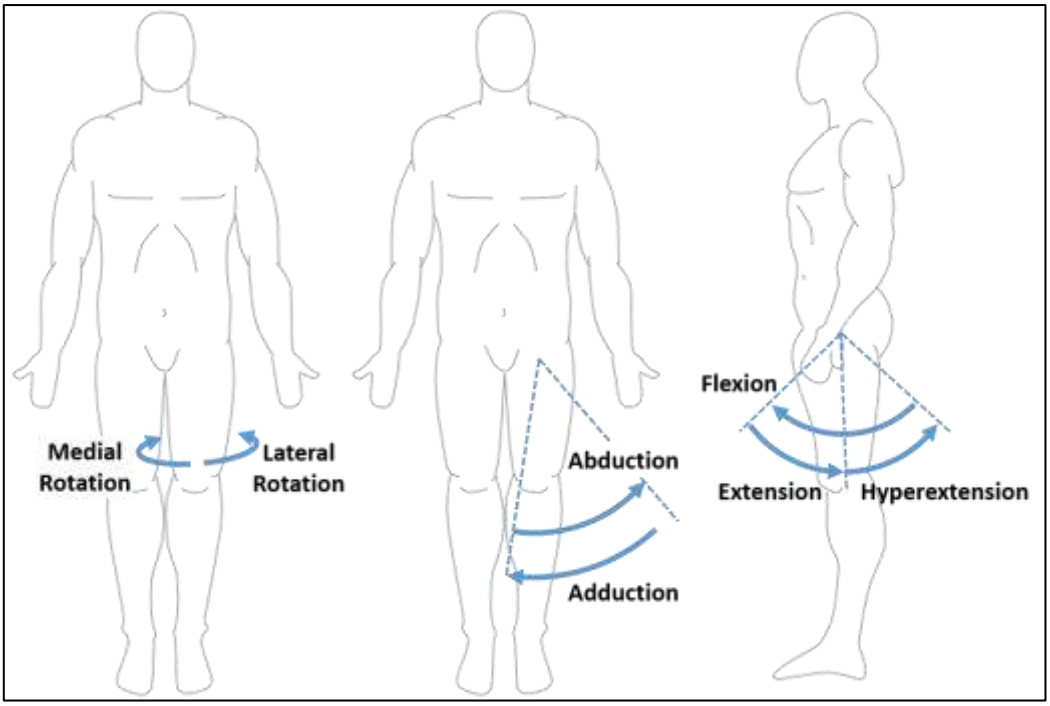

Figure 2.2: Movements of the hip. This work is a derivative of "Human body, front and side" by Nanoxyde used under CC BY-SA 3.0 licensed by A. J. FitzPatrick under CC BY-SA 3.0.

Table 2.2: Summary of hip joint movements.

Hip Movement	Description
Flexion	Moving the thigh bone (femur) forward (anteriorly)
Extension	Moving the femur in the opposite direction to flexion
Hyperextension	Moving the femur backward (posteriorly) from the anatomical position (i.e. the position depicted in Figure 2.2)
Abduction	Movement of the femur away from the midline
Adduction	Movement of the femur towards the midline
Circumduction	Movement of the distal end of the femur in a circle through a sequential combination of flexion, adduction, extension, and abduction at a joint (OpenStax, 2013e)
Medial rotation	Rotating femur about its longitudinal axis so that the anterior surface turns toward the midline
Lateral rotation	Rotating femur about its longitudinal axis so that the anterior surface turns away from the midline

2.2.2 The Hip Joint

The skeletal system of the adult human body consists of 206 individual bones; all but one of the bones articulate or form a joint with at least one other bone of the body (Tortora et al., 2016). The hip joint is the articulation of the femur (thigh bone) with the hip bone of the pelvis and is an example of a synovial joint. Figure 2.3 shows the lateral and medial views of the right hip bone and Figure 2.4 shows the anterior and posterior views of the femur bone. For the current study, the important features to note of the hip bone are the iliac crest, anterior superior iliac spine (ASIS), and the acetabulum. The important features to note of the femur are the head, neck, and greater trochanter.

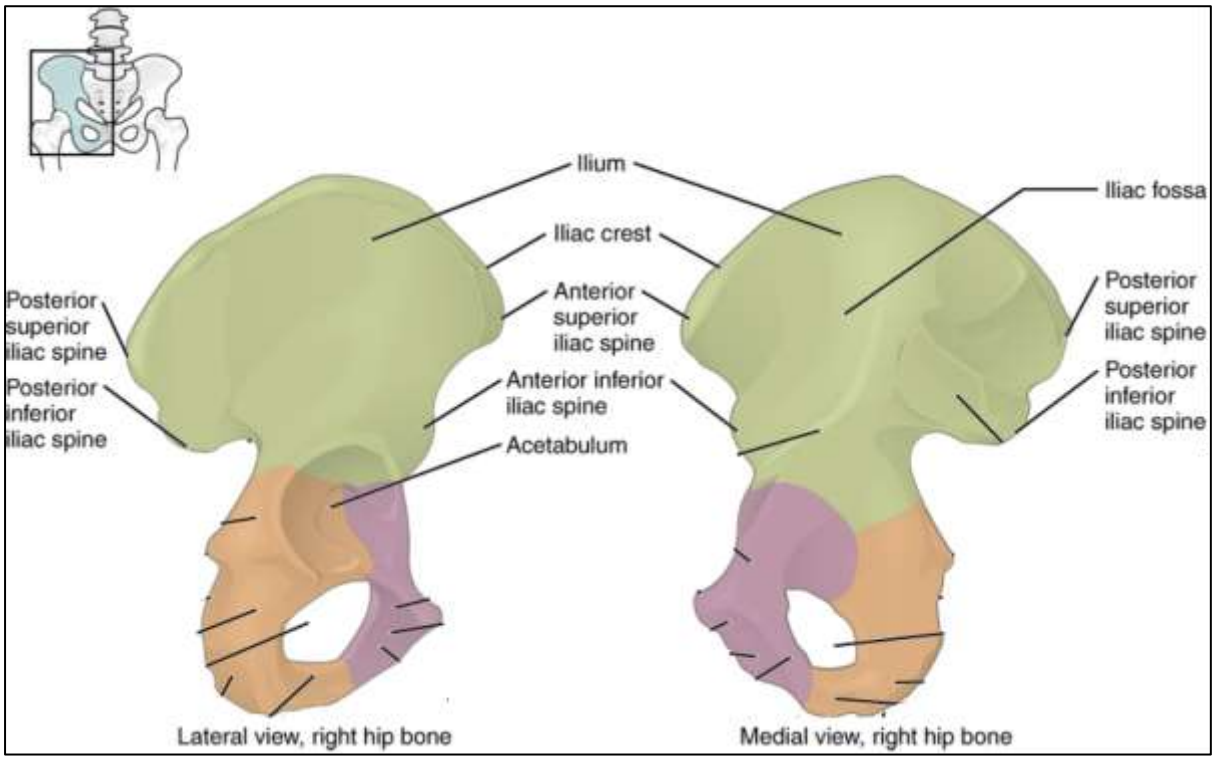

Figure 2.3: The hip bone. This work is a derivative of "The hip bone" by OpenStax (2013c, p. 3) used under CC BY 4.0 licensed by A.J. FitzPatrick under CC BY 4.0.

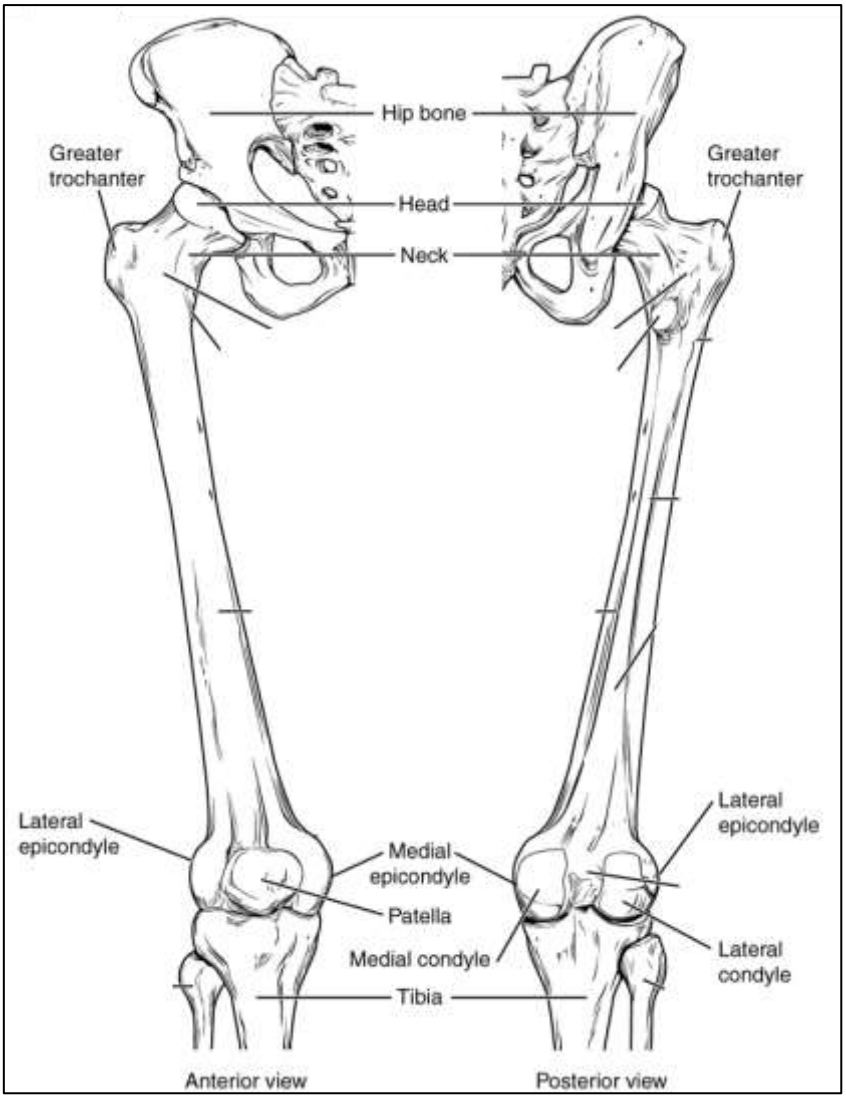

Figure 2.4: The femur bone. This work is a derivative of "The femur bone" by OpenStax (2013b, p. 2) used under CC BY 4.0 licensed by A. J. FitzPatrick under CC BY 4.0.

As identified previously, the hip joint is an example of a synovial joint. A synovial joint is defined functionally as a freely moveable joint (diarthrosis); these types of joint have characteristics that allow considerable movement between the articulating bones. There are six different structural forms of synovial joint, each form allows a different type of movement: pivot, hinge, saddle, ball-and-socket, condyloid, and plane (Tortora et al., 2016). Figure 2.5 shows a diagram with an example for each type of structural classification of synovial joint. The figure depicts that the hip joint (Figure 2.5f) is a ball-and-socket joint. A ball-and-socket joint comprises one bone with a ball-like end fitting into a cup-like socket of another bone. In the hip joint, the proximal end of the femur is ball-like and is called the femoral head. The

socket of the hip joint is a cup-like depression in the hip bone called the acetabulum. The ball-and-socket joint allows movement around three axes: flexion-extension, abduction-adduction, and rotation. (Tortora et al., 2016)

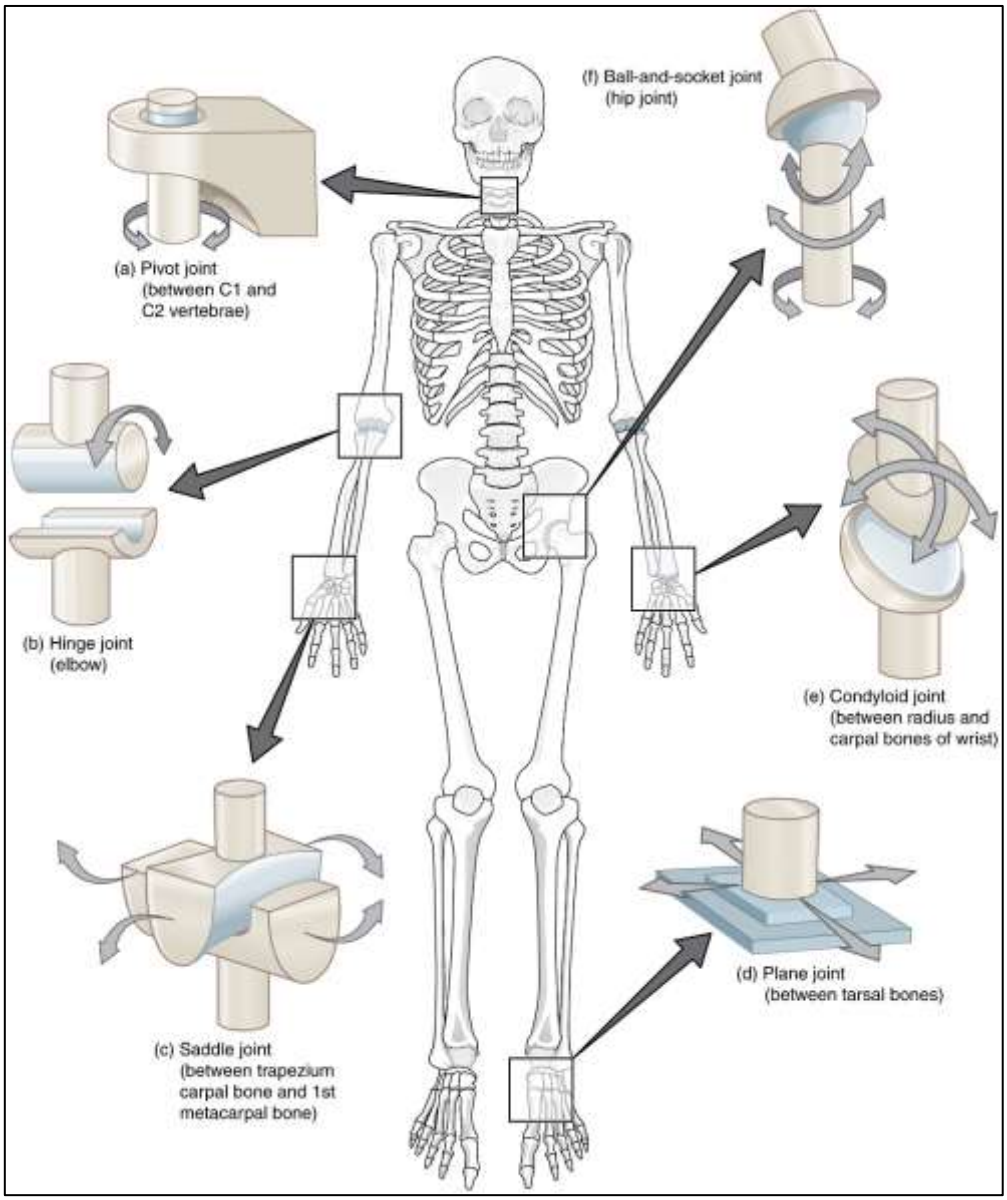

Figure 2.5: Types of synovial joints by OpenStax (2013d, p. 6) licensed under CC BY 4.0.

A synovial joint has four important characteristics that contribute to the joint being freely moveable but also remaining structurally stable: articular cartilage, an articular capsule, a synovial cavity (joint cavity), and synovial fluid. Figure 2.6a shows a cross-sectional diagram of the synovial joint of the hip identifying the important characteristics of the joint. The articulating bones of a synovial joint are each covered with a separate layer of articular

cartilage. The purpose of the articular cartilage is to reduce friction between the bones and help absorb shock; it does not bind the bones together (Tortora et al., 2016). It is this cartilage that is damaged or degraded in some way in a hip joint affected by osteoarthritis leaving the performance of the joint motion reduced and even causing pain during movement (Rabiei, 2009).

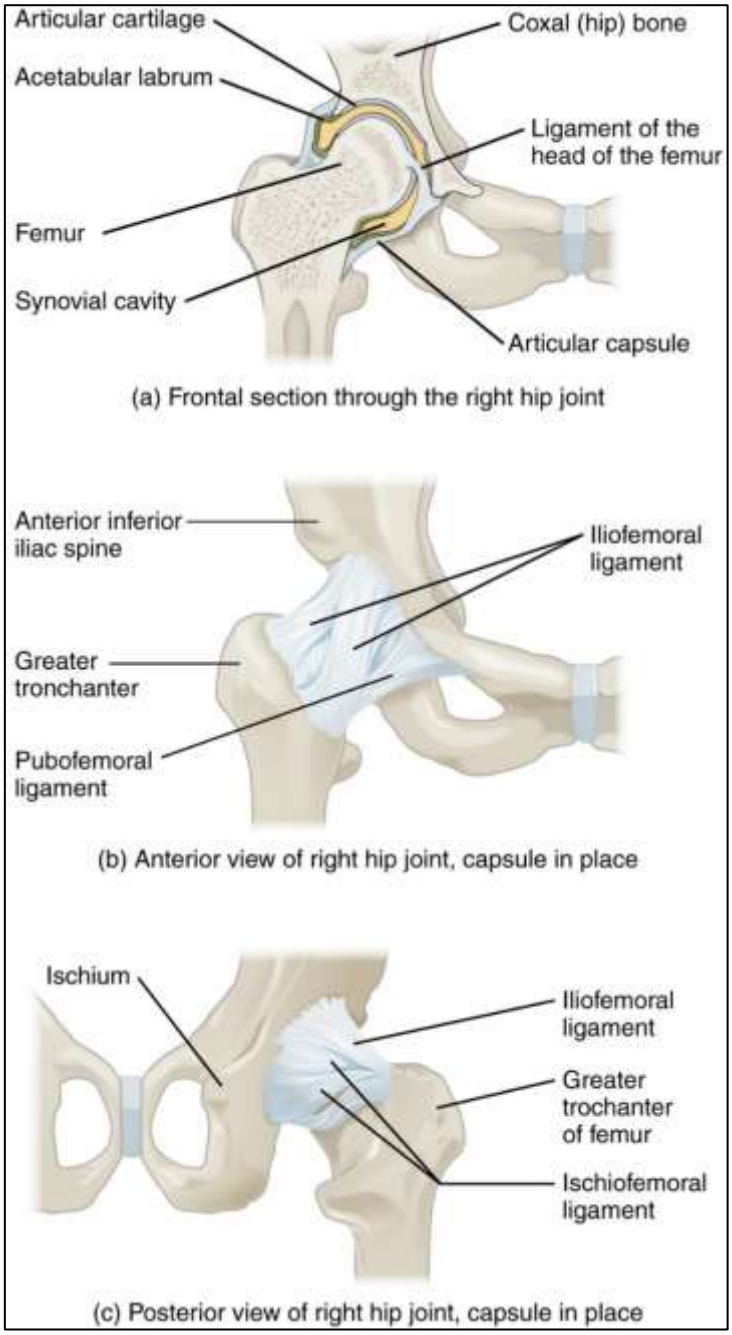

Figure 2.6: Hip joint schematic. Hip Joint by OpenStax (2013a, p. 12) licensed under CC BY 4.0.

The synovial joint is surrounded by an articular capsule which connects the bones together with a fibrous outer membrane and is reinforced by ligaments. The fibrous membrane and ligaments are flexible, allowing significant movement, and they have a high tensile strength which resists dislocation of the bones (Tortora et al., 2016). In the hip joint the fibrous membrane is reinforced by three ligaments (identified in Figure 2.6b-c): the iliofemoral, pubofemoral, and ischiofemoral ligaments. The inner membrane of the articular capsule, the synovial membrane, forms the boundary of the joint cavity between the bones and secretes synovial fluid into the cavity (Tortora et al., 2016). Synovial fluid is a viscous fluid which has a number of functions including: providing lubrication of the joint, absorbing shocks, supplying the necessary nourishment to the articular cartilage whilst also removing waste products, and removing microorganisms and debris resulting from the normal wear of the joint (Tortora et al., 2016).

2.2.3 Hip Implant Components

A THR implant takes the place of the natural femoral head and acetabulum to form a new artificial ball-and-socket joint. Therefore, a THR implant consists of both a femoral component and an acetabular component, creating the ball and the socket sides of the joint respectively. Figure 2.7a shows a diagram of the components that make up a typical THR implant prosthesis and their positioning in the body and Figure 2.7b shows an exploded view of the THR implant components. The femoral component usually consists of two modular components: a femoral stem and a femoral head. The femoral stem is inserted into a void prepared in the top of the femur after the natural femoral head has been removed. The stem is either cemented in place with bone cement or it is uncemented. An uncemented stem is press-fit into the femur and relies on bone growth into a porous surface on the stem (known as osseointegration) which consequently fixes the stem in place. The femoral head is spherical and contains a Morse taper recess which is press-fit onto a Morse taper trunnion at the neck of the femoral stem.

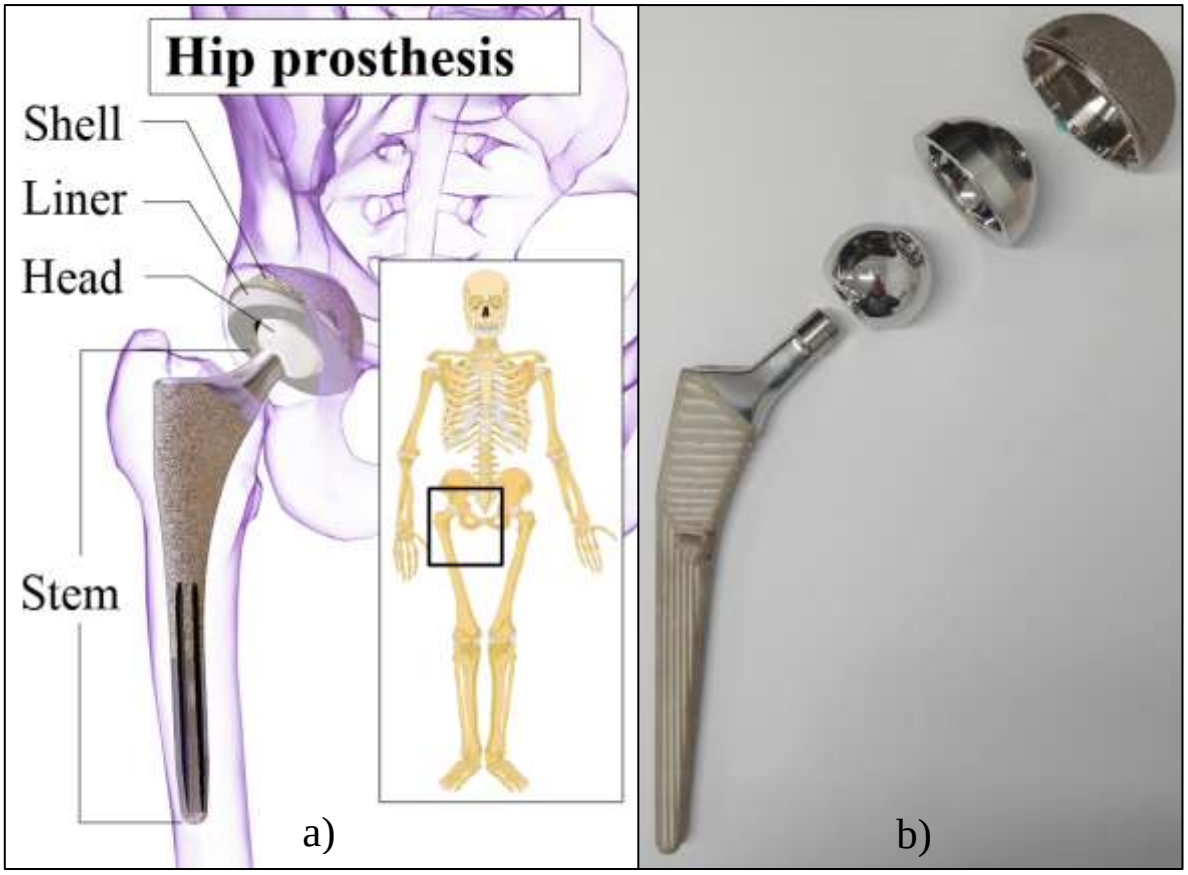

Figure 2.7: a) Components of a hip prosthesis by Mikael Häggström licensed under CC BY-SA 4.0 b) Exploded view of THR implant components.

The acetabular component also consists of two modular components: the acetabular shell or cup, and the acetabular liner. The acetabular shell is inserted into the acetabulum of the hip bone after the acetabulum has been prepared by removing the articular cartilage and shaped to accept the acetabular shell. Similar to the femoral stem, the fixation of the acetabular shell is either cemented or uncemented. The acetabular shell will usually contain a Morse taper socket of which a Morse taper on the outer surface of the acetabular liner is press-fit into. The acetabular liner contains a spherical depression in which the femoral head mates. Hence, the main articulating or bearing interface of the implant is between the femoral head and the acetabular liner. In some cases, when a cemented acetabular shell is used, an acetabular liner is not required because the shell itself contains the spherical depression that accepts the femoral head.

The acetabular shell and femoral stem are typically made from titanium or cobalt-chromium alloys due to the excellent biocompatibility of these materials. The acetabular liner and femoral head are usually made from one of either polyethylene (PE), alumina (ceramic), or cobalt-chromium-molybdenum alloy (metal) which creates a number of possible bearing surface combinations (Hutchings, 2003; Rabiei, 2009). Some of the bearing surface combinations used in practise are known as hard-on-soft and consist of a metal femoral head on a PE acetabular liner (MoP) or a ceramic femoral head on a PE acetabular liner (CoP). Other combinations are known as hard-on-hard and consist of either a ceramic femoral head on a ceramic acetabular liner (CoC), a ceramic femoral head on a metal acetabular liner (CoM), or a metal femoral head on a metal acetabular liner (MoM). The latter of these bearing surface combinations, MoM, is becoming more uncommon as there is increasing evidence of undesirable effects on patient health due to metal-ion debris (Browne et al., 2010). Conversely, advances in material design has produced highly cross-linked polyethylene bearing surfaces which has been shown to have significantly lower wear rates than conventional polyethylene and hence hard-on-soft bearing surfaces using highly cross-linked polyethylene are considered the gold standard (Liu et al., 2015; Pivec et al., 2012).

2.2.4 Hip Implant Failure Modes

The most common cause of late failure of THR has been aseptic loosening as a result of particle debris from the articulating surfaces of the replacement (Abu-Amer et al., 2007; CIHI, 2015; Delaunay et al., 2013; Garellick et al., 2015; NZOA, 2015; Pivec et al., 2012; Ulrich et al., 2008). Aseptic loosening is a loosening of an implant component's bond with the bone in the absence of any infection. The New Zealand Orthopaedic Association (2017) reported that in New Zealand 33% of THR revisions in 2016 were for aseptic loosening (acetabular 17.8%, femoral 15.2%). The implication for a patient with aseptic loosening is that loading forces

cannot be effectively transferred and micromotion will occur between the implant and surrounding bone leading to production of wear and debris particles (Drees et al., 2007). Hence, loosening is likely to cause pain and increase the potential for breakdown of the bone or periprosthetic bone fractures (fractures of bone in close proximity to an implant) and ultimately total implant failure. Surgical intervention is necessary to treat aseptic loosening and will involve the exchange of either the femoral stem or acetabular cup, or both of these components. Any surgical intervention which involves the exchange of components interfacing with the bone will be more time consuming, more expensive, and increase the recovery time of the patient. Early detection and intervention of osteolysis and implant loosening can mean less invasive surgery and is therefore beneficial in terms of both the patient outcomes and use of health resources.

Implant instability and dislocations are another major cause for revision of THR implants (CIHI, 2015; Garellick et al., 2015; NZOA, 2015; Pivec et al., 2012). Instability in a joint is when there is greater than normal movement of the joint in any direction causing dissociation of the articular surfaces. Implant malpositioning and abductor deficiency (weakness of abductor muscles supporting the hip joint) are the two most common factors influencing THR instability and recurrent dislocations. Other factors include component sizing, patient compliance to rehabilitation strategies, patient age, and component wear among others (Parvizi et al., 2008; Pivec et al., 2012). The vast majority of dislocations after THR surgery are solitary occurrences and can be treated non-operatively. However, recurrent dislocations usually require surgical intervention. Surgical intervention in these cases generally involves the exchange of the implant acetabular liner with a constrained version or the use of a larger diameter femoral head (Parvizi et al., 2008). In 2016, 16.9% of THR revisions in New Zealand were for implant dislocation (NZOA, 2017).

Another reason for revision of THR components is implant squeaking in patients who have implants with hard-on-hard bearing surfaces such as ceramic-on-ceramic or metal-on-metal. Squeaking typically begins fourteen to twenty-six months after surgery (Keurentjes et al., 2008; Restrepo et al., 2008; Walter et al., 2007) and can be disconcerting and embarrassing for patients, impeding their quality of life (Jarrett et al., 2009; Owen, D. et al., 2014). The origin of squeaking is controversial and has been repeatedly reported as multifactorial; dependent upon implant and surgical factors. Implant factors relate to friction at the bearing surface (Brockett et al., 2013; Restrepo et al., 2008; Taylor et al., 2007; Walter et al., 2011; Walter et al., 2007; Walter et al., 2008; Weiss et al., 2010), femoral stem design (Restrepo et al., 2010), third body wear (Brockett et al., 2013; Chevillotte et al., 2010) and micro-separation between the ceramic head and liner (Nevelos et al., 2000) whereas surgical factors include acetabular orientation (Sexton et al., 2011; Walter et al., 2007; Walter et al., 2008), neck length (Keurentjes et al., 2008) and rim impingement (Ecker et al., 2008; Parvizi et al., 2011). Some researchers believe patient characteristics also influence squeaking as those who are younger, taller and with a higher body mass index experience increased rates (Sexton et al., 2011; Walter et al., 2007), suggesting that implants in these patients would be subject to greater forces and stresses (Walter et al., 2007).

A recent meta-analysis found the incidence of squeaking in ceramic-on-ceramic THRs to be 4.2%, of which the incidence of subsequent revision for squeaking was 0.2% (Owen, D. H. et al., 2014). Even though the literature shows that there are much more common reasons for implant revision than squeaking, squeaking has the unique characteristic of being identifiable to the naked ear. Hence, the phenomenon of implant squeaking offers a unique opportunity to investigate implant AEs, with some certainty that the emissions are induced by the presence of the implant, to aid in implant condition monitoring and diagnostics.

Periprosthetic bone fracture, especially of the femur, is also a relatively common cause of THR revision. Fractures in close proximity to the implant can cause a range of complications which would generally lead to joint instabilities and/or implant loosening and hence early intervention is necessary to prevent total failure of the implant. The New Zealand Orthopaedic Association (2017) reported that in New Zealand 14.4% of THR revisions in 2016 were due to a fractured femur. Pain and infection are also two other common reasons for revision of THR implants. In 2016, 13.5% of THR revisions in New Zealand were indicated to be due to pain and 13.1% due to infection (NZOA, 2017). It is important to note that the statistics given for the revision reasons may be inflated due to the fact that it is common for multiple revision reasons to be specified for a given implant and all reasons are included in the analysis (NZOA,2017).

One final failure mode that is worth mentioning is fracture of the implant components themselves. Due to advances in material design and manufacturing technologies, fracture of modern implant components is uncommon (e.g. third generation ceramic femoral heads have an approximate fracture rate of 0.004% as opposed to a rate of 13.4% for early generation ceramics (Sivananthan et al., 2014)). Nevertheless, revisions for fractured implants still occur in THR patients with early generation ceramic components or occasionally in patients with implants that have been in service for many years.

2.3 Acoustic Emissions in Medicine

2.3.1 Acoustic Emissions in the Human Body

One of the most recognisable symbols identifying a doctor is the presence of a stethoscope around their neck. The modern stethoscope has played an important and incredibly useful role in patient diagnoses ever since it's invention by the French physician, Rene Theophile Hyacinthe Laënnec, in 1819 (Roguin, 2006). Laënnec discovered that the use of a stethoscope

(mediate auscultation) meant that heart sounds (or any sounds in the chest cavity) could be heard with a greater clarity than by placing his ear directly to the chest (immediate auscultation) (David & Dumitrascu, 2017; Roguin, 2006). Immediate auscultation was a diagnostic technique known in ancient times (for example, it was used by figures such as Hippocrates in ancient Greece) but it was well known that the technique was less than ideal for reasons such as modesty, bad hygiene, and attenuation in obese patients (David & Dumitrascu, 2017; Roguin, 2006). The modern stethoscope has contributed greatly to the field of medicine (and continues to do so) in terms of advancements in the knowledge and accuracy of heart and lung disease diagnoses through the study of sounds emanating from the human body (David & Dumitrascu, 2017; Kaniusas, 2015; Roguin, 2006).

Internal body sounds are generally initiated from mechanical vibrations of tissues, blood, muscles, and air in the airways and are particularly prevalent in the thorax of the body (Kaniusas, 2015). Sounds generated from the heart are arguably the most recognised body sounds and are commonly used to aid in diagnosis of heart problems due to the repeatability observed in normal heart sounds. Lung sounds are another prominent body sound which allow information on the condition of the airways to be observed. Lung sounds are more variable over time and have been explained by Laënnec as easier to distinguish than to describe (Kaniusas, 2015). Lung sounds tend to be continuous in nature and normal lung sounds generally do not have recognisable features in the time domain (Kaniusas, 2015). Abnormal lung sounds can be continuous and have a seemingly musical character, or they can be discontinuous and have an explosive or bubbling character usually heard as crackling (Kaniusas, 2015). Heart sounds have been observed to have a frequency range up to 400 Hz, with lung and snoring sounds having frequencies from approximately 100 Hz up to 1000 Hz and 2000 Hz respectively (Kaniusas, 2015).

Another common internal body sound that can occur is clicking as a result of the movement of joints. Non-symptomatic clicking is common in the general population (Guillin et al., 2012). These non-symptomatic clicks can be induced from tendons, ligaments, muscles, or even bone impinging with, and then subsequently moving rapidly over, bony prominences ("Any questions?," 1969; Chandran Suja & Barakat, 2018; Guillin et al., 2012; Kawchuk et al., 2015). When tendons, ligaments, or muscles are tight they are more likely to impinge on each other or on bones during joint movement. When the increasing tension overcomes the friction of the impingement, the tendon, ligament, or muscle can rapidly move past the impingement to relieve tension and consequently releases vibrational energy that can be heard as a click.

A further mechanism that is thought to induce non-symptomatic clicking of joints is the rapid separation of joint surfaces, known as articular release (Protopapas & Cymet, 2002). Clicking of this type is commonly identified as the noise that occurs when “cracking” one’s knuckles or back. There is still much debate as to the exact cause of clicking resulting from articular release but it is generally thought to be a result of either the formation of a gas bubble (cavitation) in the joint cavity or the subsequent rapid collapse of the cavitation (Castellanos & Axelrod, 1990; Chandran Suja & Barakat, 2018; Kawchuk et al., 2015). Regardless of the exact mechanism, the clicking from articular release is generally not believed to be detrimental to the health of the joint (Castellanos & Axelrod, 1990; Protopapas & Cymet, 2002).

Symptomatic clicking of joints is commonly a result of problems in the direct vicinity of the joint such as foreign or loose bodies, labral tears, cartilage defects, or torn ligaments and usually has associated pain (Guillin et al., 2012; Lewis, 2010; Protopapas & Cymet, 2002; White et al., 2004). Continual clicking as a result of the aforementioned problems can result in further damage to the tissue concerned or surrounding tissue. Hence, if pain is associated with joint

clicking then it is advised that clinical investigation occurs to properly diagnose any abnormalities (Guillin et al., 2012).

It is evident from the literature that the range of sounds that emanate from various internal regions of the human body are widely used to aid clinical examinations and diagnoses. The study of these sounds by medical practitioners is common due to the ease at which they can be identified and monitored. The sounds propagate to the skin surface and can therefore be non-invasively detected with an instrument as simple as a stethoscope. Additionally, sometimes the body sounds have enough intensity to propagate past the skin surface and all the way to an external observers ears without the need for amplification. Hence, it would be unwise to ignore such obvious and easily monitored sounds when assessing the health of patients. It is apparent from the literature that many sounds that are generated from the human body are considered normal and can, in many cases, indicate good health. Conversely, some body sounds offer no diagnostic assistance as they are seemingly occurring randomly or they are not yet well understood. Nevertheless, it is obvious that internal body sounds are common and the study of these sounds plays an important role in the identification of various medical conditions.

2.3.2 Acoustic Emission Technique in Orthopaedics

Primitive AE analysis on human joints, primarily the knee joint, has been carried out by medical practitioners since the early 1900's using stethoscopes and without the use of signal capture and processing techniques. These early experiments provided promising evidence of the usefulness of AE monitoring in joint injury characterisation and early diagnosis. The development of data acquisition and signal processing techniques over the past half century have also allowed significant advancement in the analysis of AE signals (Glaser, 2008). A study conducted over four years from 1985 analysed the main frequency content of AEs from

a number of total hip replacements in patients in the clinical environment (*in-vivo*) and concluded that there was good correlation between the main frequency content and implant fixation condition (Gao et al., 1990). Other *in-vivo* studies in the 1980's concluded that an increase in the number of AE events was related to clinical symptoms and radiolucency at the bone/cement interface of total knee replacements (Kohn, 1995).

Gao et al. (1990) used two AE transducers, with a frequency range up to 1 kHz, attached to the skin of the patient during walking, sitting down, and standing up motions. The study found that the main frequency content of the AEs from patients with apparently loose implants were either below 50 Hz or above 500 Hz and those for patients showing no signs of loose components were between 200 and 350 Hz. Additionally, the study found that no AEs were detected from participants with natural joints while the occurrence of AEs from THR patients had good reproducibility. The study also indicated the AE technique could detect very early stage loosening that X-ray inspection could not despite using only the main frequency component as a diagnostic metric of loosening. Abnormal AEs were detected in some patients where X-rays showed no signs of abnormalities. These patients were re-monitored after a time of rest-cure and were found to have an improved clinical condition.

The conclusions from the study by Gao et al. (1990) show great promise for the diagnostic potential of the AE technique in orthopaedics. However, the study limited its AE analysis to frequencies below 1 kHz and subsequent studies have shown AEs from THR implants to have more notable content occur with frequencies higher than 1 kHz (Currier et al., 2010; Khan-Edmundson et al., 2012; Sariali et al., 2012; Walter et al., 2008). A study, carried out at the University of Canterbury, used AE monitoring to investigate the attenuation characteristics of the soft tissue surrounding THRs (Khan-Edmundson et al., 2012). It was found that the maximum vibration frequencies that were present at the skin surface were approximately 20

kHz and that there was noticeable attenuation of signals above 10 kHz. It should be noted however that this study was not comprehensive and made broad conclusions only and therefore its conclusions should be considered with caution.

A study at the University of Tennessee combined AE monitoring of THRs with vibration and video fluoroscopy gait analysis (Glaser et al., 2008). Video fluoroscopy is a technique which records a series of successive X-rays to allow viewing of moving structures radiographically. Glaser et al. (2008) conducted the study in order to correlate AEs and vibrations with *in-vivo* joint kinematics through the video fluoroscopy. The study investigated 24 patients, with various implant bearing interface combinations. It was found that separation and subsequent meeting of the implant's main bearing interface correlated well with AE generation and that the various interface materials produced characteristic AE profiles (Glaser, 2008; Glaser et al., 2008, 2010). Since the study focussed primarily on the relationship between AEs and the separation of the implant bearing surfaces, it did consequently relate the implant separation over time to stages of the gait cycle. Hence, the time domain signals of the observed AEs were indirectly related to the stages of the gait cycle. However, in terms of indicating how consistent any trends were between the AE signals and gait cycle stages, Glaser et al. (2008) only went so far as to state that a "knocking sound was observed when the femoral head contacted the acetabulum" and that "squeaking and rolling sounds were detected when intensive back-and-forth movement of the femoral head occurred." This conclusion regarding "knocking," "squeaking," and "rolling" sounds indicates that there is likely a valuable relationship between implant AEs and stages of the gait cycle but this was not the general conclusion of the study. Furthermore, the study did not attempt to relate the observed AEs to the condition of the participants' implants since the participants all were clinically evaluated to have excellently functioning implants.

Many of the *in-vivo* studies to date have contributed important findings with regard to the validity of AE monitoring of joint replacements but few have assessed AEs from large and representative cohorts. It is evident from published studies that the AE monitoring technique has a place in orthopaedics; it has been shown that *in-vivo* AE signals from different THR patients show varied behaviours but seem to correlate with various clinical diagnoses and implant motions. However, many of the recent studies which assess AEs of THR patients and relate them to clinical diagnoses use only a small number of patients, sometimes less than ten, and hence the validity of these relations is difficult to confirm.

There has also been a significant amount of *in-vitro* studies on the AEs of joint replacements. This research has been done in an attempt to gain further insight into the AE source as well as AE characteristics of implant loosening, loadings, and failures. *In-vitro* studies carried out at the University of Southampton have investigated bone cement damage, stress development during cement cure, and failure processes of carbon fibre reinforced plastic (CFRP) prosthesis components (Browne et al., 2010; Browne et al., 2005). It was found from these studies that AE monitoring could; locate and monitor damage in bone cement, monitor crack development in curing cement, and assist the understanding of failure in CFRP components. Another *in-vitro* study investigated the effect on the incidence of squeaking of various component designs as well as bearing clearances and loading (Hothan et al., 2011). It was found that component design had an important effect on the squeaking characteristics and that load increases caused corresponding increases in the squeaking frequencies. Further *in-vitro* literature investigates such things as; AE characteristics of bone (Kohn, 1995), AEs of implant loosening in porcine limb specimens (Ruther et al., 2013), and diagnosis of loosened THRs using induced vibrations in an *in-vitro* hip model (Rieger et al., 2013).

Hothan et al. (2011) conducted *in-vitro* AE testing on un-lubricated THR components to determine the influence of cup and stem designs on squeaking incidence and characteristics. Eigenfrequencies of the assembled femoral components were found to be good predictors of squeaking frequencies and higher axial loads caused an increase in these squeaking frequencies. Brockett et al. (2013) investigated the potential causes of squeaking of hard-on-hard hip bearings *in-vitro* and concluded that squeaking correlates well with increased friction at the bearing interface. However, the study used a simplified model of the gait cycle and short test periods on unworn implants and so the effect of long term wear was not tested. Additionally, the study was not extended to include any *in-vivo* testing.

Rashid and Pullin (2014) critically review AE testing research in a variety of Orthopaedic applications. They found well-conducted *in-vitro* studies that validate the use of AEs for early detection of aseptic loosening in femoral components of THR (Davies et al., 1996; Mavrogordato et al., 2011; Sugiyama et al., 1989). Piriou et al. (2016) investigated a finite element analysis (FEA) model that can reproduce squeaking frequencies seen *in-vivo* and verified it using an *in-vitro* hip simulator. The study found that modal frequencies from the FEA and the *in-vitro* AE frequency content were consistent with each other. When soft tissues were included in FEA analysis the observed frequencies lowered to be a closer match to squeaking frequencies recorded *in-vivo*.

2.4 Technical Background Information

2.4.1 Traditional Damage Monitoring Using Acoustic Emissions

The AE monitoring technique is well established in the non-destructive testing (NDT) of engineering components and materials as it is effective at detecting fracture and fatigue in a range of materials and applications. Early application of the AE technique had a focus on

detecting and monitoring crack growth in engineering structures (Eitzen & Wadley, 1984; Scruby, 1987). Crack growth causes a sudden change in stress and a subsequent displacement within the material resulting in an elastic (stress) wave radiating through the material away from the crack; commonly known as a primary AE event (Ohtsu et al., 2016; Scruby, 1987). Additionally, AEs can be caused by friction between any surfaces (e.g. friction between the crack faces, known as crack fretting). When a system has more kinetic energy than the associated friction can dissipate, vibration instabilities are normally introduced which can result in the production of AEs (Akay, 2002). AEs induced by friction processes are commonly known as secondary AE events and do not necessarily indicate crack growth (Ohtsu et al., 2016; Scruby, 1987).

The presence of secondary AE events has allowed the AE technique to be extended for use in monitoring and control of production processes and in the testing of engineering materials. The technique has the ability to continually monitor the normal AEs (generally secondary AE events) from manufacturing processes (such as welding, cutting, forming, turning, milling, grinding, and drilling) and detect when abnormal emissions occur (Baranov et al., 2007; Ohtsu et al., 2016; Scruby, 1987). Abnormal AEs from manufacturing processes are generally identified as a significant primary AE event which is uncharacteristic of the normal process or a secondary AE event where a sustained change has occurred in what is considered the normal properties of the AEs (for example, the noticeable presence of a new frequency in the typical frequency profile of the process). The presence of abnormal AEs can indicate damaged or loose tools and components, and specific locations can usually be identified if multiple sensors are used (Gholizadeh et al., 2015; Ohtsu et al., 2016; Scruby, 1987). AEs can also distinguish between specific failure mechanisms of various materials and help advance the understanding of material deformation and fracture (Gholizadeh et al., 2015; Scruby, 1987).

A review of the AE technique by Gholizadeh et al. (2015) identified the value of using the technique in the structural health monitoring of civil structures through the detection of cracking early. AE monitoring can be employed to detect the initial formation of cracks and can follow their progression. Therefore, the technique provides real-time information about the integrity of a structure allowing an immediate response to be made. AE monitoring also provides information across a wide range of materials and hence has been applied to monitor the structural integrity of concrete, steel, or wooden civil structures, in addition to pressure vessels and piping systems (Gholizadeh et al., 2015). The AE technique has also been employed extensively in aerospace applications because of its ability to effectively monitor damage and failure of composite materials in real-time. Specifically, the AE technique has been used in aircraft damage monitoring to detect fatigue cracks, barely visible impact damage, delamination, and corrosion (Staszewski et al., 2004). The technique has also demonstrated the ability to differentiate between different mechanisms of composite failure such as matrix cracking, fibre-matrix debonding, and delamination (Gholizadeh et al., 2015).

2.4.2 Acoustic Emission Propagation

As stated in Section 1.3 an AE is a mechanical elastic wave that propagates through a medium. This propagation occurs due to the fact that when a particle is excited by an AE it vibrates in its location but dynamically interacts with adjacent particles (Ohtsu et al., 2016). The particle interaction results in energy being transferred between particles and hence the wave propagates. The elastic wave is classified from considering the direction of propagation in relation to the direction of the particle vibration (Kohn, 1995; Ohtsu et al., 2016).

A primary wave (P-wave) is a longitudinal wave where the particles vibrate in the direction of the wave propagation. A secondary wave (S-wave) is a transverse wave (sometimes called a

shear wave) where the particles vibrate in the direction perpendicular to the wave propagation. Figure 2.8 shows a schematic of transverse and longitudinal waves as a visual aid. Other waves that occur in a solid are Rayleigh waves and Lamb waves (Kohn, 1995; Ohtsu et al., 2016; Scruby, 1987). These other waves occur when a P- or S-wave reach the boundary or surface of the propagation medium. Rayleigh waves result in an elliptical motion of the particles near the surface of the medium as shown in Figure 2.9. Lamb waves occur in thin specimens and cause the whole specimen to vibrate in either a symmetric or anti-symmetric mode as shown in Figure 2.10.

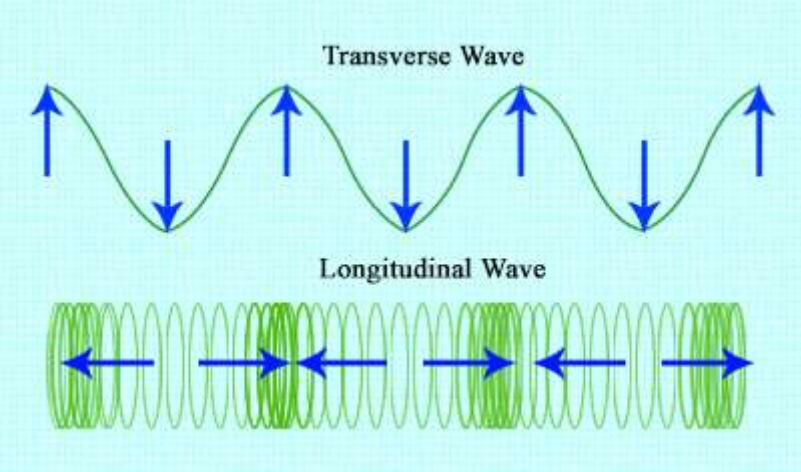

Figure 2.8: Schematic of transverse and longitudinal waves by Zappys Technology Solutions licensed under CC BY 2.0.

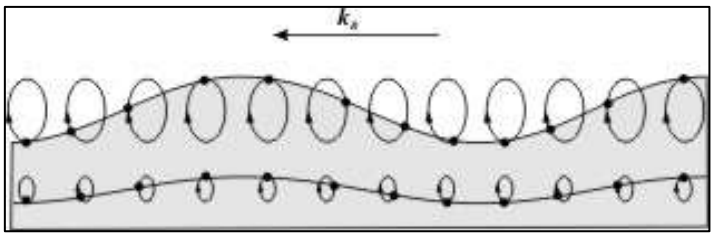

Figure 2.9: Rayleigh wave propagation by MPasternak at Polish Wikipedia licensed under CC BY 2.5.

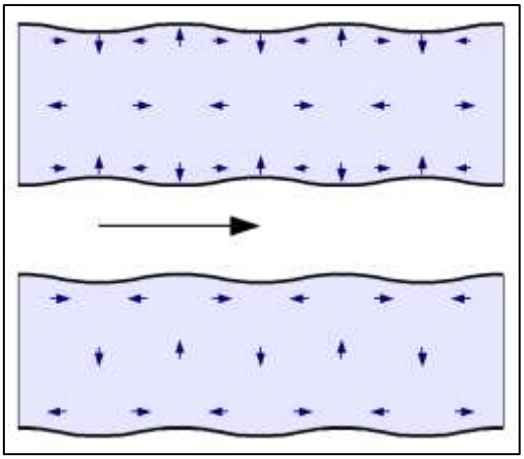

Figure 2.10: Lamb wave schematic showing a symmetric mode vibration (top) and an anti-symmetric mode vibration (bottom). Lamb wave schematic by Michael Schmid licensed under CC-BY-SA-3.0 and GFDL.

P-waves propagate with a greater velocity than S-waves and the wave velocity is determined by the elastic modulus, density, and Poisson's ratio of the propagation medium (Ohtsu et al., 2016). A Rayleigh wave has a velocity approximately 10 percent slower than a transverse wave in the same medium and the velocity of a Lamb wave is dependent on the dimensions of the medium, and the frequency and vibration mode of the wave (Ohtsu et al., 2016). The differences in wave speed can influence how an AE will be detected by a sensor. In general, an AE sensor would detect longitudinal waves first, followed by transverse waves. However, when waves (longitudinal or transverse) meet a boundary in the medium, additional longitudinal and transverse waves will be generated as a result of refractions and reflections (Kohn, 1995; Ohtsu et al., 2016; Scruby, 1987). These additional waves can result in an AE sensor detecting reflected or refracted longitudinal waves before the originally generated transverse wave is detected. Furthermore, the occurrence of reflected and refracted waves is likely to mean many vibrations detected by an acoustic sensor are the result of wave superposition and hence the detected AE signal is likely to include distortions.

Attenuation of acoustic waves also occurs during propagation and will influence the vibrations detected by an acoustic sensor. Viscous damping in a material causes temporal attenuation of vibrating particles and the greater the distance of a particle from the AE source the greater the

damping of the vibration amplitude (Ohtsu et al., 2016). Additionally, the attenuation of AEs increases noticeably as the vibration frequency increases (Khan-Edmundson et al., 2012; Ohtsu et al., 2016). As a result of AE attenuation, the extent of the aforementioned wave superposition and associated distortion may be reduced when the waves reach the acoustic sensor thus distortions may be less significant. Furthermore, when acoustic sensors are on the surface of an object, Rayleigh waves tend to dominate the detected acoustic waveforms due to the fact that they propagate in two dimensions, as opposed to three-dimensions, hence Rayleigh waves experience less attenuation (Scruby, 1987).

The propagation characteristics outlined above are important to remember when considering the propagation of AEs in the THR system. If the joint components of a THR are the source of an AE then the acoustic waves must travel through a number of various different substances to reach an AE sensor on the skin surface. Gao et al. (1990) presented a transfer function (TF) for the propagation of an acoustic wave from the source to a sensor located at either the posterior-superior iliac spine or femoral lateral condyle. The overall TF for a system is the product of the specific TFs of each different material along the propagation path. The TF presented by Gao et al. (1990) is likely the most simple that could be applied to the THR system yet it included nine individual TFs for various components along the propagation path. Some of the components that make up the propagation path of AEs in the THR system are: the prosthesis, bone, numerous types of soft tissues (such as fibrous tissues, tendons, ligaments, muscles, and skin), and even acoustic coupling gel that is used between the skin and the sensor membrane.

Given that the overall TF is a product, it is obvious that any output signal detected at a sensor on the skin surface is highly influenced by the components that constitute the specific propagation path of the acoustic wave. Determining the specific transfer function for THR AEs detected on the skin surface is outside the scope of the current thesis. However, it is important

to bear in mind the greater complexity of the wave propagation paths in biological systems compared to those of AEs in industrial applications where commonly only one or two different materials may be involved. The increased complexity of the biological system is likely to affect comparisons of AEs across different sensor locations and severely limit the ability to triangulate the location of AE sources. Nevertheless, differences in the bone and soft tissues of THR patients with and without implant complications will consequently result in differences in the overall TFs. Hence, differences in detected AEs should be observed between patients with and without implant complications.

2.4.3 Acoustic Emission Parameters

An AE wave or signal that is detected by an acoustic sensor can be characterised by a number of parameters that describe the detected waveform. There are parameters based on the intensity and frequency of the waveform (such as peak amplitude, root-mean-squared (RMS) amplitude, acoustic counts, and main frequency content) and the shape of the waveform (such as duration, rise time, and the measured area under the rectified signal envelope (MARSE)). Generally, any acoustic signal measured by an acoustic sensor has a base level of noise greater than zero. Therefore, an arbitrary amplitude threshold greater than the noise level is used to determine the presence of an AE event. An AE event is usually defined as an AE pulse that exceeds the given threshold and then decays back to the noise level (Baranov et al., 2007; Ohtsu et al., 2016; Rao, 1990; Scruby, 1987; Shark et al., 2010). Consequently, parameters such as duration, rise time, MARSE, and acoustic counts are highly dependent on the threshold value. The AE activity of an AE signal is defined as the total number of discrete AE pulses (i.e. AE events) per unit time.

As stated above, an AE event is identified when the acoustic signal exceeds the arbitrarily chosen amplitude threshold. Figure 2.11 shows a typical waveform with annotations identifying the common signal parameters which are generally defined as follows:

- Duration – the time span from the first crossing of the threshold to the last crossing of the threshold before the signal amplitude returns to the base noise level.
- Peak amplitude – the maximum absolute voltage value within the duration of the AE.
- Rise time – time span from the first crossing of the threshold to time of the peak amplitude.
- RMS amplitude – the voltage defined by $A_{RMS} = \sqrt{\left[\sum_{i=1}^{N} x_i^2\right]/N}$ where x_i is the voltage of the i^{th} sample from the first threshold crossing of the AE event and N is the total number of samples within the duration of the AE event.
- MARSE – the integral of the rectified signal envelope over the duration of the AE event.
- AE counts – the total number of times a signal crosses the amplitude threshold in the positive voltage direction.
- Main frequency content – the frequency corresponding to the greatest magnitude in the frequency domain of the AE event (i.e. the frequency that corresponds to the peak magnitude of the Fourier transform of the AE signal).

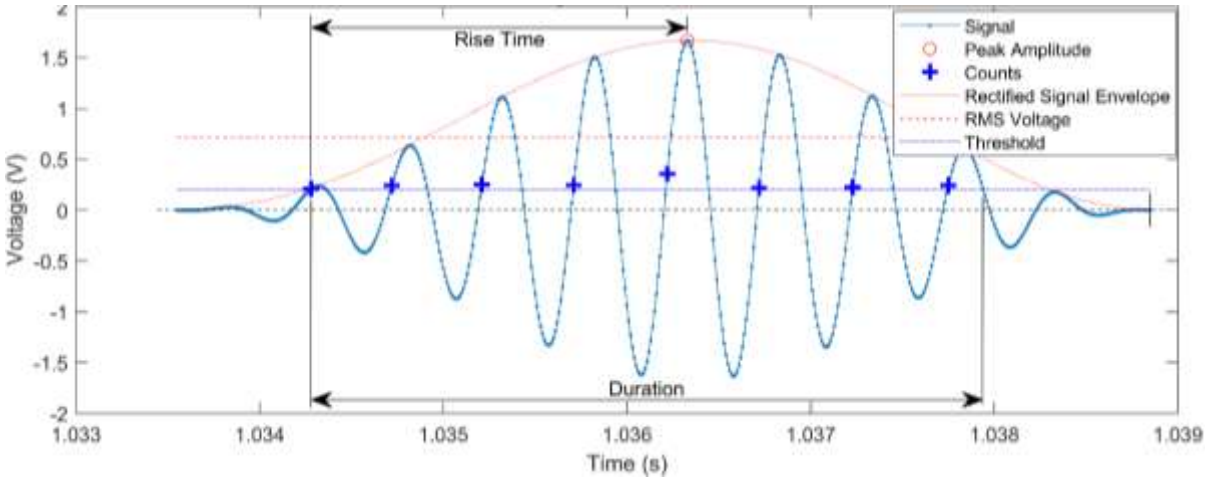

Figure 2.11: Typical waveform and its basic characteristics.

The duration, peak amplitude, rise time, RMS amplitude, and main frequency content parameters are important to note. The aforementioned five parameters are important to keep in mind for the following chapters of this thesis since they will form the basis of the analyses that

are presented in Chapter 4 and Chapter 5. The MARSE and AE counts parameters will not be used since the equivalent information will be available from the RMS amplitudes and main frequency content respectively. The MARSE of a signal will be correlated with the RMS amplitude as both measures are related to the energy of the signal. The AE counts of a signal is related to the average frequency content since the AE counts is a measure of the number of cycles per total duration of the event.

2.4.4 Acoustic Emission Event Detection

Recorded AE signals from THR patients generally contain large regions of signal with amplitudes that do not exceed the background noise. Consequently, the AE recordings contain only small regions of interest (i.e. AE events) within much longer signals. Therefore, AE event detection was employed in the analyses presented in this thesis and methods for event detection are reviewed here. The most common technique used to detect events in voltage signals is to identify when the signal exceeds a set voltage threshold. In this technique, an arbitrary voltage threshold is chosen that distinguishes the signal from noise. An event is generally defined as the region between the first and last crossing of the threshold before the signal amplitude returns to the noise level (Baranov et al., 2007; Ohtsu et al., 2016; Rao, 1990; Scruby, 1987; Shark et al., 2010). A voltage threshold for event detection has good discrimination of events for signals with a high signal-to-noise ratio. However, the method does not perform well when the signals are weak as it can be difficult to determine an appropriate voltage threshold value which accurately detects events. The thresholding method can be modified to use a variance signal where the variance of a sliding window over the signal is used instead of the original voltage signal. Applying the thresholding method to the variance signal can result in better discrimination of events since random fluctuations in the noise level are smoothed out allowing a lower threshold value to be used to detect events.

An alternative event detection method investigated by Rodgers et al. (2017) used the statistical variation in sound intensity. Given the AE data collection hardware and methodology of Rodgers et al. (2017), the statistical variation in sound intensity method assumes the sound intensity is proportional to the voltage squared. Therefore, given that the sound intensity of ambient noise remains relatively constant, AE events can be defined as the outliers in the sound intensity when plotted over time. The outliers/events were determined by splitting the sound intensity values into 0.1 second time windows and checking if the maximum value for a window exceeded three standard deviations above the mean for the entire signal. To further refine the detection of events to capture low energy events, an iterative process was used where the standard deviation was recalculated after removing any time windows with detected events from the signal and checking for new outliers. The iterative process was carried out until no new outliers were detected. Rodgers et al. (2017) concluded that the sound intensity method performed well but further research was required to properly assess the effectiveness of the method. One key disadvantage with this method is that the AE events were effectively defined as constant duration windows in time. Despite removing the need for an arbitrary voltage threshold, this definition of an event limits the ability to effectively compare AE event parameters since most of the parameter definitions of Section 2.4.3 are dependent on using a voltage threshold.

Wavelet transform event detection has been used by a number of studies in literature with varying success (Ahadi & Bakhtiar, 2010; He et al., 2010; Kadambe & Boudreauxbartels, 1992; Ng & Qi, 2001; Rodgers et al., 2017). The wavelet method utilises a short oscillating signal known as a mother wavelet which has a non-zero amplitude for a finite time period. A series of additional wavelets are created by shifting and scaling the mother wavelet with respect to time. The correlation between each created wavelet in turn and the AE signal can then be calculated and output as a wavelet coefficient. The larger a wavelet coefficient value, the better

the associated wavelet matches the AE signal. Therefore, if the shape of the mother wavelet closely matches the waveform of specific AE events, large wavelet coefficient values will indicate the presence of an AE event at the corresponding time in the signal. However, if the shape of the mother wavelet does not match the waveform of an event well the wavelet coefficients may remain small and no event would be able to be identified. Consequently, the wavelet event detection method is limited by the ability to select a mother wavelet that closely matches the shape of the signal that needs to be detected. Given that the shape of THR AE event signals is not initially known and multiple types of AE event waveforms may be present, multiple wavelet analyses would need to be conducted to both determine the mother wavelets to be used and to detect all types of events.

A further event detection method investigated by Rodgers et al. (2017) used the root-mean-squared (RMS) amplitude of the AE signals. This method created an RMS amplitude signal by calculating the RMS amplitude of a small signal window as the window was moved over the entirety of the AE signal. AE events were determined by applying an amplitude threshold to the RMS amplitude signal and identifying the regions where the RMS amplitudes exceeded the threshold. Different methods for determining the threshold value were investigated such as user-selected, using the RMS value during noise-only regions, and using the mean value of the RMS amplitude signal. Rodgers et al. (2017) concluded that the RMS amplitude event detection method with the threshold set by using the mean value of the RMS amplitude signal was less robust than using a user-selected threshold but it performed well for analysing large patient datasets since it was efficient to implement.

Teague et al. (2016) used a signal envelope detection algorithm to perform event detection of knee joint sounds. The method used a short time Fourier transform (STFT) with two millisecond time windows and defined the envelope as the logarithmic amplitude of the STFT

coefficients for each time window in turn. A moving mean and standard deviation of the envelope was calculated using a 1000 sample window. The occurrence of an event was determined by thresholding the envelope using a value of 3.3 standard deviations from the moving mean. The method performed well at locating the occurrence of knee clicks. However, the method stopped short of defining event start and end points.

Another method for event detection is the use of pattern recognition and machine learning algorithms (Grzeszick et al., 2017; Rao, 1990). In general, pattern recognition and machine learning algorithms that are applied for AE event detection rely on using short time windows of the AE signals. Various signal time and frequency parameters, such as those presented in Section 2.4.3 (excluding duration dependent parameters), can be calculated for each short time window. The signal parameters as a function of time can then be passed through machine learning algorithms, such as clustering or neural networks, to identify patterns in the data. If repeated signal patterns are identified by the algorithms they are likely to indicate the occurrence of similar AE events. However, the accuracy of machine learning algorithms is called into question if the nature of the input data does not effectively capture the parameters that characterise different AE events. This is difficult to determine without processing a large number of input parameters and visually verifying their results. Furthermore, any supervised machine learning techniques would require large training data sets of identified AE events which is impractical to obtain for the THR AE dataset.

Of the event detection methods outlined above, all methods have their advantages and disadvantages but no single method has an obvious advantage over the others when considering the scope of this thesis. With respect to the detection of AE events without having the subjectivity of using arbitrary thresholds, machine learning algorithms are an obvious choice. However, the development and implementation of a machine learning algorithm for AE event

detection is outside the scope of this thesis. While AE event detection forms an important part of this research, emphasis is placed on implementing a simple AE event detection algorithm that allows an initial investigation into the basic time and frequency domain characteristics of the regions of interest of the AE signals. For this reason, a simple event detection algorithm based on thresholding of the variance of the AE signal will be employed in the analysis of this thesis.

2.5 Summary

This chapter presented relevant background information for three key topics: anatomical terminology and the THR implant; literature of AEs in medicine; and technical aspects of the AE technique and AE signals. Section 2.2.1 presented important terminology regarding the planes of the human body (Figure 2.1a), body directions (Figure 2.1b and Table 2.1), and hip joint motions (Figure 2.2 and Table 2.2) that is used throughout the remainder of this thesis. An introduction to the anatomy of the natural hip joint was presented accompanied by an explanation of the THR implant and its components. The four main components of a THR (femoral stem, femoral head, acetabular liner, and acetabular shell) were presented in Figure 2.7. Additionally, the common failure modes of the THR implant (aseptic loosening, dislocation, excessive noise, periprosthetic bone fracture, pain, infection, and implant fracture) and their corresponding prevalence were outlined in Section 2.2.4.

Section 2.3 provided an overview of literature covering AEs in the human body and specifically AEs in the orthopaedic field. The literature demonstrated that, experimental results from *in-vitro* implant tests have provided further understanding of the AE behaviour of joint replacements. However, *in-vitro* experimental results from the literature have not yet directly contributed to the *in-vivo* diagnosis of total hip or knee replacement complications. No clear

diagnostic method or clear indication of signal characteristics that can reliably predict specific failure modes currently exists. Some promising initial results have been observed in the literature but more research is needed into larger patient groups and more in-depth signal processing analyses to better understand how AEs relate to clinical outcomes. Furthermore, the inclusion of gait analysis and other sensing methods has shown promise of being able to provide additional insight into *in-vivo* implant mechanics. However, the literature concerning THR AEs combined with gait analysis is extremely sparse.

The final sections of this chapter presented technical background of the AE monitoring technique, AE signal characteristics, and AE event detection methods. It was shown from literature that AE monitoring is usually employed in traditional damage monitoring situations to detect large discrete AEs that indicate crack growth or identify abnormal emissions in processes where the normal AEs are predictable and well-defined. Additionally, in many traditional damage monitoring applications the AEs can be detected from the same material that they originated in. However, it was explained that the propagation characteristics are much more complex for AEs travelling from inside the human body to the skin surface since they potentially pass through many different materials (such as the implant components, bone, fibrous tissues, tendons, ligaments, muscles, and skin). Therefore, the increased complexity of the biological system is likely to affect the characteristics of the AEs. Nevertheless, differences in the bone and soft tissues of THR patients with and without implant complications is hypothesised to result in detectable differences in AE characteristics. The hypothesis of detectable differences from AEs with different implant complications is the foundation of the subsequent research presented in this thesis.

Chapter 3 Data Collection Protocols

3.1 Introduction

This chapter presents the details of the AE monitoring hardware and formalises the data collection protocols for the *in-vivo* and *in-vitro* AE testing. First, the specific hardware and associated parameters of the AE monitoring device is explained. A statement of ethical approval is then presented. Following that, the formalised data collection protocol for the *in-vivo* AE monitoring is explained. Finally, the *in-vitro* test apparatus and associated testing protocol is presented. A condensed version of this chapter forms the methods section of the published work FitzPatrick et al. (2017b).

3.2 Acoustic Emission Device

The AE detection device that was used for data collection utilised four passive ultrasonic receivers/sensors (Prowave 328ER250) each with a resonant frequency of 32.8 kHz. The signals from the sensors were amplified by a circuit mounted within each sensor housing in order to prevent additional noise adding to the low voltage signal during transmission to the data acquisition (DAQ) system. The sensors were connected to a National Instruments CompacDAQ through a NI-9222 analog module. The CompacDAQ was connected via USB to a computer running LabVIEW software where the voltage signals from the four sensors were recorded simultaneously at 100 kHz. The 100 kHz sampling rate provided a maximum practical frequency of 50 kHz by the Nyquist theorem which was well in excess of the typical frequencies observed during prior *in-vivo* and *in-vitro* implant testing (Khan-Edmundson et al., 2012).

The sensors were cylindrical in shape and had a diaphragm diameter of 25.1 mm. Each sensor was mounted in a custom built housing to allow the attachment of the signal amplification circuit and to increase the robustness of the device. The sensor housings were constructed from acetal plastic and were cylindrical in shape with an overall diameter of 40 mm and a height of 33 mm. Figure 3.1a shows an image of one of the sensors assembled inside its housing with the diaphragm of the sensor visible. The design of the sensor housing meant the diaphragm of the sensor was recessed into the housing by approximately 1.5 mm. The recessing of the sensor meant reducing the chance of damaging the diaphragm and restricting vibration through direct contact with hard objects such as the implant components during *in-vitro* testing for example. Instead, during all testing scenarios, standard ultrasound transmission gel was used to cover the sensor diaphragms and create a reliable transmission path for AEs to the sensors from either the skin surface (during *in-vivo* testing) or the implant components (during *in-vitro* testing). The combination of the recess and ultrasonic gel also avoided undesirable DC voltage offsets which would have resulted from constant contact pressure of a hard object against the sensor diaphragm. Figure 3.1b shows the acoustic sensors, organised in an array ready for *in-vivo* testing, and the CompacDAQ with analog module. Also identified in the Figure is a signal breakout box which was used to create an input point for the sensors' power source.

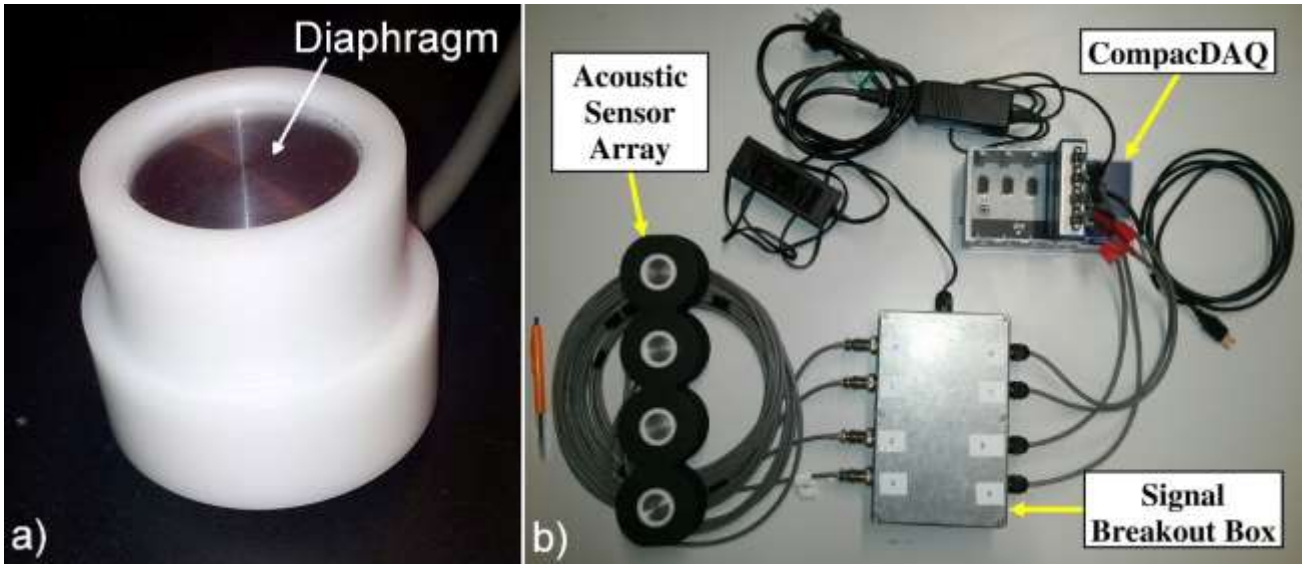

Figure 3.1: a) Assembled acoustic emission sensor and housing. Note the sensor diaphragm (visible inside the white housing) has a diameter of 25.1 mm. b) Image of data acquisition hardware which consists of four ultrasonic sensors arranged in an array, signal breakout box for external power input, and National Instruments CompacDAQ (A ball-point pen of length 140mm has been included at the left of the image for scale).

3.3 Ethical Approval & Participant Recruitment

Ethical approval for *in-vivo* AE data collection was granted from the New Zealand Upper South A regional ethics committee under approval number URA/10/11/075. The ethics approval allowed the recruitment and AE monitoring of persons who had THR implants along with a control group of 20 people with healthy, natural hips (allowing data for 40 control hips). AE monitoring was nominally performed at the Canterbury Orthopaedic and Bone Research Association (COBRA) unit at Burwood Hospital in Christchurch but alternate locations could be used subject to additional locality agreements.

Following the ethics approval, a cohort of THR patients were recruited for *in-vivo* monitoring. The recruitment criteria included patients scheduled for revision surgery in Christchurch with identified implant problems, such as loosening or audible squeaking, as well as those identified through routine questionnaires or consultations. The patients had a range of age, implant type, and implant conditions. Additionally, a second cohort of participants who had healthy natural hip joints with no history of hip joint trauma or injury were recruited to establish a control

database of AEs from natural hip joints. Prior to commencement of the AE monitoring of a participant, an informed consent process was carried out to ensure each participant understood the purpose of the study and what the AE monitoring procedure involved. If the participant agreed to partake in the study the appropriate consent form was completed.

3.4 *In-Vivo* Data Collection

3.4.1 Sensor Configuration & Attachment

For *in-vivo* testing, the four sensors were arranged in a linear array using a closed cell foam pad to position the sensors. The foam pad had four evenly spaced holes for the sensors to locate into where the distance between the centres of adjacent holes was 70 mm. The holes in the foam pad had diameters approximately 2 mm smaller than the sensor housings to provide an interference fit with the sensors. The interference fit meant that the sensors were held in position tightly and their positions relative to each other would not change. Figure 3.2 shows a photograph of the sensor pad array with the 70 mm spacing identified.

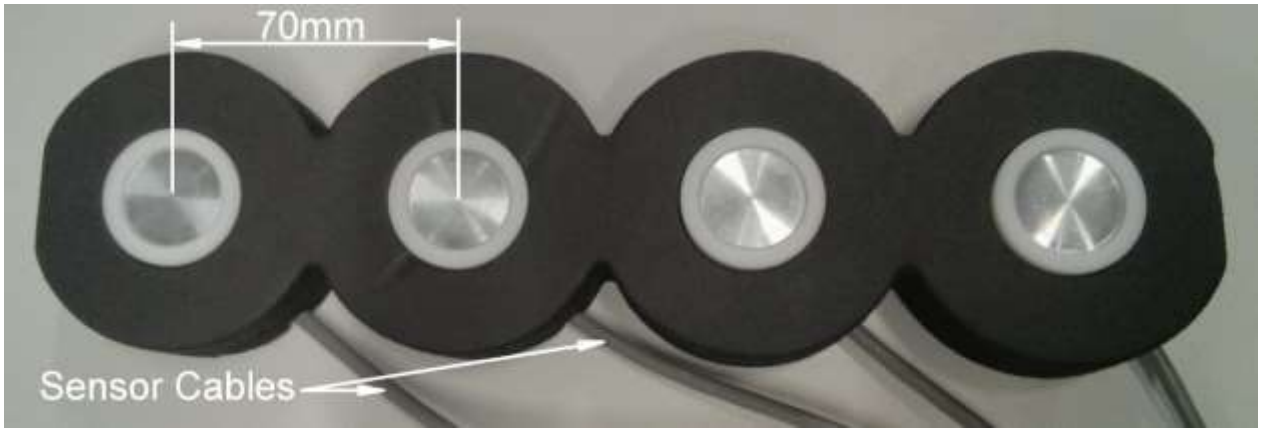

Figure 3.2: Sensor pad array used for in-vivo AE monitoring.

The sensor array was placed against the skin surface of the participant from the iliac crest to the upper-femur. The positioning of the sensors relative to the *in-situ* implant are illustrated in Figure 3.3. Note that, while the photograph in Figure 3.3 shows the sensor array on the outside of the clothing, this is only to demonstrate the position on the patient. During actual in-vivo

testing the sensor pad is directly against the skin of the patient not outside the clothing as shown in Figure 3.3. Each sensor was numbered in ascending order from one through four starting from the highest placed sensor on the participant and concluding with the lowest placed sensor (see Figure 3.3). This numbering scheme of the sensors remained consistent for all *in-vivo* testing so as to allow comparisons of AEs from particular sensors across different participants. As mentioned in Section 3.1, ultrasound transmission gel was used on each sensor to ensure a reliable sound transmission path by filling the space between the skin and the sensor diaphragms. The sensor array was held in place using crepe bandages around the waist of the participant. This method of fastening the sensor array to the participant was simple, efficient, and reliably ensured continuous contact of the sensors with the skin during testing.

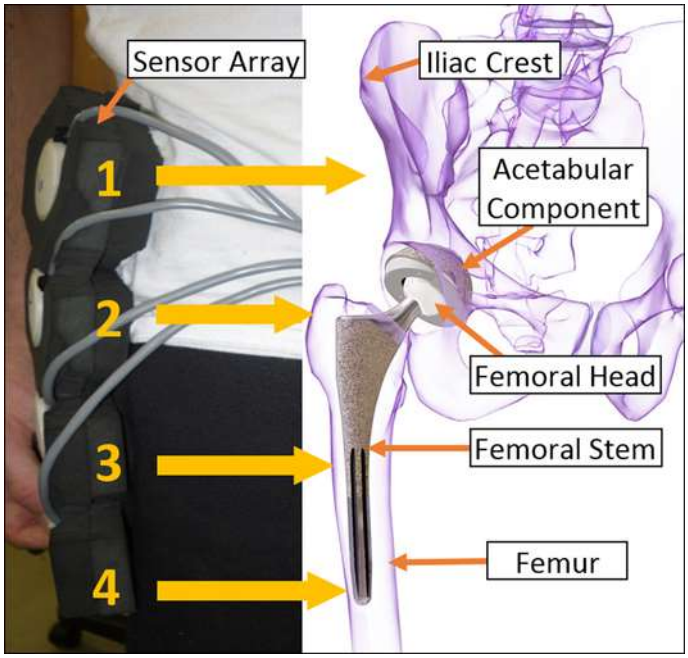

Figure 3.3: Schematic showing the positioning of the four ultrasonic sensors relative to the in-situ implant and surrounding skeletal system (Note that during actual in-vivo testing the sensor pad is against the skin of the patient not above the clothes as shown in the Figure here). This work is a derivative of "Total Hip Replacement" by BruceBlaus used under CC-BY-SA 4.0 licensed by A. J. FitzPatrick under CC-BY-SA 4.0.

3.4.2 Test Motions

Following attachment of the sensor array, each participant was asked to perform a number of basic motions. The main purpose of the test motions was to subject the hip joint of the

participant to a number of dynamic everyday ranges of motion and loadings whilst concurrently measuring any AEs that occurred from or near the joint. The participants were free to perform each prescribed motions in a way that best suited their health and mobility. The full range of motion for different patients varied due to their mobility. Participants were not placed at risk by pushing them outside their comfortable range of motion. Therefore, the specific execution of each motion varied from one participant to the next. However, the movement of the hip joint still remained broadly similar between participants. Additionally, if a participant was not able to, or did not wish to, perform a particular motion then that motion was left out of their test procedure all together.

The first of these motions was a squatting motion where the participant performed a standard squat from a standing position. The participant typically held onto the back of a chair in front of them help maintain their balance during the squat in the interest of safety. The illustration of Figure 3.4 shows the basic squatting motion, the position of the person portrayed in the figure is considered a shallow squat as opposed to a deep squat where the hips would drop below the knees. In general most of the participants performed very shallow to shallow squats since they moved only as far as they felt comfortable. Enforcing a deep squat in patients with identified problems that were scheduled for revision surgery would place them at high risk of pain and/or potential injury and/or dislocation of their implant. While it introduces a limitation to this study, patient safety had to take priority over obtaining a larger range of motion. A typical AE recording for the squatting motion began with the participant standing and recorded a single squat and finished once the participant was standing upright again. Each participant performed approximately six squats.

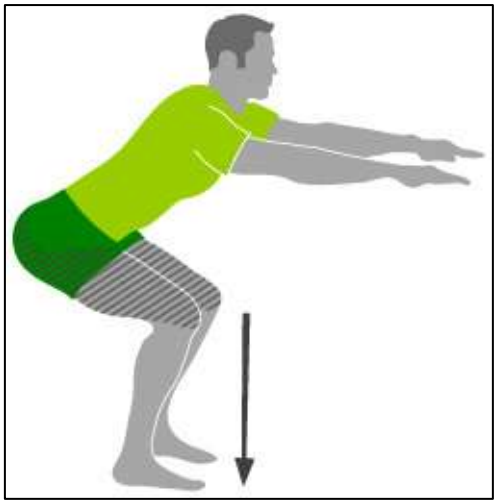

Figure 3.4: Illustration of squatting motion. Kniebeuge by Marcel Kollmar licensed under CC-BY-SA 3.0 DE.

The second motion that the participants were asked to perform was standing from, and lowering/sitting into, a seated position. The participants began seated on a chair and moved to a standing position in their own time. Once standing, the participants moved back into a seated position on the chair. Some participants could not perform these motions easily unassisted and therefore used a walking stick or the back of another chair in front of them to help stabilise and pull themselves up and to steady themselves as they sat back down. Typically, each standing and sitting motion were recorded as separate AE recordings. However, on some occasions both a standing and sitting motion were recorded as part of the same AE recording. Generally, each participant repeated the stand and sit motion at least three times.

The third motion that the participants were asked to perform was a stair ascent and descent. For practical reasons a step-ladder was used in place of a staircase to perform this motion. The participants were asked to walk up the step ladder one step at a time as far as they felt comfortable and to use the step-ladder handles in front of them for balance and support. In general, each participant had both feet on one step before they continued onto the next. Typically, one AE recording captured the ascent steps of the participant and a second recording captured the descent steps. However, on some occasions the ascent and descent steps were contained within one recording. Generally, at least three repetitions of the stair ascent and

descent motion was performed by each participant. The participants were not asked to lead their steps with a specific leg and hence a mixture of leading with the implanted leg and the opposite leg occurred. However, some participants voluntarily alternated their leading leg for each repetition of the motion. This combination provides both loaded and unloaded motion of the monitored hip joint.

The fourth motion was basic walking. The participants were asked to simply move around the testing room by walking as normally as possible. The tethered AE sensors and the size of the testing room meant that the participants were limited to an area of approximately six square metres to walk. The limited walking area meant that only three to five paces could be performed in a straight line and hence the walking motion generally included participants walking in arcs as they turned around. The AE recordings for the walking motion always included multiple paces and typically recorded for three seconds or more. Generally, three or four AE recordings of three to six seconds or more were recorded for the walking motion.

The final prescribed motion that the participants were asked to perform was heel strikes. From a normal standing position the participant was asked to repeatedly strike their heel against the ground by lifting and lowering their entire leg in a stomp-like motion. The purpose of the heel strikes was to determine how any heel impact was transmitted to the implant and/or surrounding tissue. The heel strikes were carried out using only the leg that contained the hip implant being monitored. The participants struck the ground with only as much force as they felt was comfortable. Hence, the force at which the heel struck the ground varied considerably between some participants. The AE recordings for the heel striking motion included multiple repeated heel strikes and typically recorded for three seconds or more. On average, three AE recordings of three or more seconds were recorded for the heel strike motion.

The participants were also given the opportunity to perform any other motion which they believed induced unique acoustic events from their prosthesis, such as squeaking. These motions ranged from bending over at the hips, to leg extensions from a seated position, to lunges. The additional motions were repeated at least three times and, depending on the motion, the AE recordings captured either a single repetition or multiple repetitions.

3.5 *In-Vitro* Data Collection

3.5.1 Overview

A portion of the patients whose implants were monitored *in-vivo* went on to have subsequent revision surgery of all or some of their implant components. A number of these revision patients gave consent for their implant components that were removed during revision surgery to be used for *in-vitro* testing. The *in-vitro* testing involved manipulating the retrieved implant components in the laboratory environment to replicate the typical joint motions of flexion/extension and rotation. Any *in-vitro* AEs generated during these motions would be compared to the *in-vivo* AEs from the same implant components. A further purpose of the *in-vitro* testing was to investigate and compare the AEs generated from the different component interfaces of the implant.

Two different methods of implant manipulation were used for the *in-vitro* testing; robot arm manipulation and manual manipulation. The robot arm manipulation was employed to provide repeatable and reliable imitation of basic hip joint flexion/extension and rotation motions of the implant. Therefore, the robot arm manipulations were primarily concerned with main bearing interface interactions of the implants and the associated AEs. Whereas, the purpose of the manual manipulation was to investigate AEs generated from the trunnion Morse taper interactions of the femoral components that the robot arm test apparatus was unable to perform.

It should also be noted that in many revision cases only the bearing components (femoral head and acetabular liner) of the implant were replaced during revision surgery, i.e. the femoral stem and acetabular cup remained in the patient and were fitted with a new femoral head and acetabular liner respectively. Hence, in these cases, the patient's acetabular cup and femoral stem could not be tested *in-vitro*. Therefore, when only the femoral head and acetabular liner were available for testing, these retrieved bearing surfaces were mated with alternative, but closely matched, unused acetabular cup and femoral stem components for the *in-vitro* testing.

3.5.2 Robot Arm Test Apparatus

The robot arm test apparatus utilised a 4-axis Epson RC20 robot arm to manipulate the retrieved implant components. The robot arm end effector was a cylindrical shaft with the shaft axis aligned in the vertical direction. The end effector was able to be moved in the three Cartesian coordinate directions x, y, and z where the x and y axes were in the horizontal plane and the z axis was in the vertical plane. Additionally, the end effector could be rotated about the vertical shaft axis (z-axis). A coordinate system is shown in-set on Figure 3.5.

Figure 3.5 shows an image of the robot arm with an implant ready for *in-vitro* testing. It can be seen from Figure 3.5 that the components were arranged in the *in-vitro* apparatus in an inverted orientation; i.e. the implant was orientated with the femoral stem pointing upwards. The inverted orientation was used to allow manipulation of the femoral stem with the robot end effector, while the acetabular cup remained stationary, was considered to be the simplest method of reproducing realistic joint motions with the robot arm.

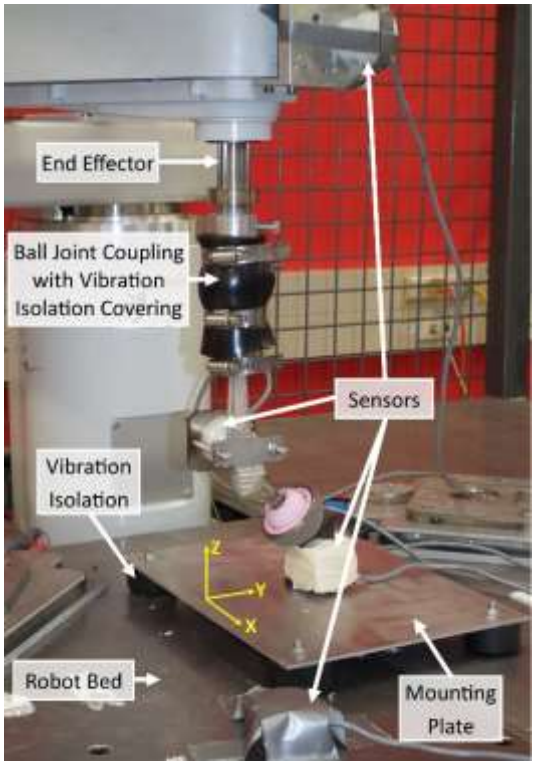

Figure 3.5: Robot arm in-vitro test apparatus showing implant ready for testing.

It can also been seen from Figure 3.5 that an additional coupling was used to connect the end effector to the femoral stem of the implant. The additional coupling included a ball joint connection (known as a rose bearing) which allowed the movement of the robot end effector to change the orientation of the femoral stem while the femoral head remained in contact with the acetabular liner. Figure 3.6a shows the *in-vitro* test setup with the end effector positioned so the femoral stem is at the maximum flexion position. It can be observed from the figure that the additional coupling is necessary as the end effector must remain vertical and the femoral head must remain in contact with the acetabular liner concurrently. As a result, there is a need to allow relative rotation between the robot end effector and the femoral stem, which is provided by the rose bearing.

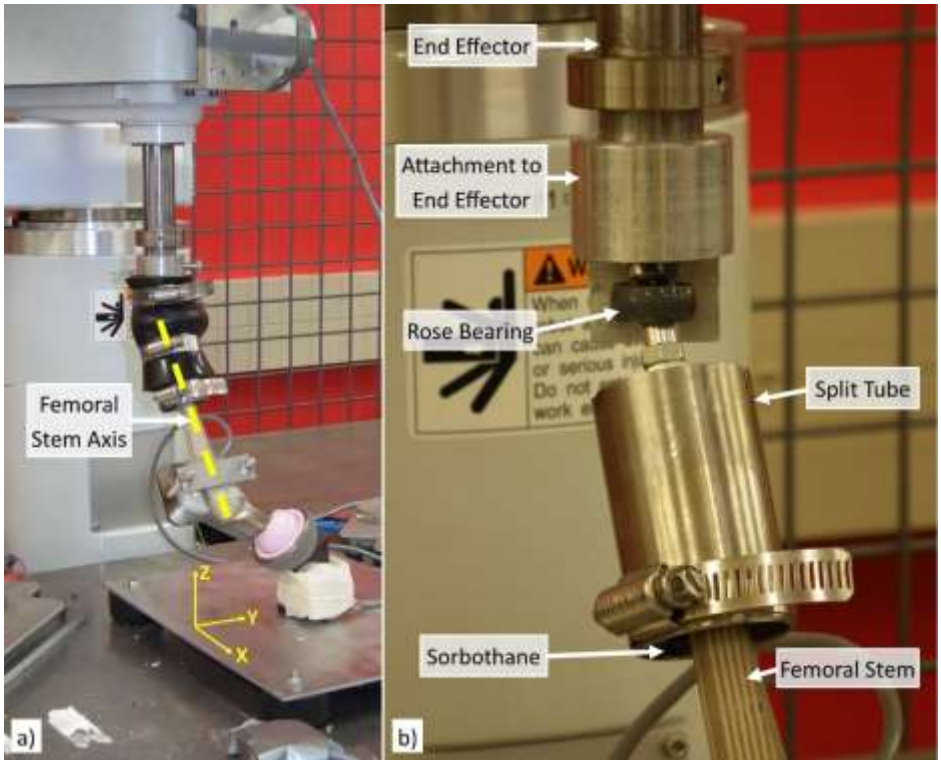

Figure 3.6: a) Robot arm in-vitro test apparatus showing implant in maximum flexion position. b) Close-up of ball joint coupling with Sorbothane covering removed.

The testing apparatus employed Sorbothane (an industrial grade vibration isolation material) to allow force transfer but isolate the implant from any unwanted external vibrations produced by the robot arm actuators. Initial testing showed that the vibrations produced by the robot spanned the same range of frequency as the implant emissions and that this prevented simple low-pass or high-pass filtering to remove the robot vibrations. To isolate the acetabular components from vibrations propagating through the robot bed, a mounting plate with four Sorbothane mounts was screwed to the robot bed. The acetabular cup was secured to a steel mounting block which was attached to the centre of the mounting plate. To isolate the femoral components from vibrations propagating through the robot arm, Sorbothane was incorporated into the coupling between the end effector and the femoral stem.

A close-up view of the coupling with the Sorbothane covering removed is shown in Figure 3.6b. A layer of 5 mm thick Sorbothane was wrapped around the end of the femoral stem which was inserted into a tube with a 3 mm wide split running the length of the tube. Once the femoral

stem and Sorbothane were in place inside the tube, a hose clamp was used to compress the tube until the split closed up and the femoral stem was held in place. Another layer of Sorbothane was then clamped to the outside of the split tube and to the attachment above the ball joint. The purpose of the outer layer of Sorbothane was to stop the femoral stem freely rotating due to the rose bearing connection and to transmit torque from the end effector to the femoral stem.

The *in-vitro* testing utilised the same sensing hardware as the *in-vivo* patient testing but alternative sensor locations were required due to the different setup involved. As shown in Figure 3.5a, one of the ultrasonic receivers was attached directly to the femoral stem and a second was attached to a steel block mounted directly below the acetabular cup. The two additional sensors were mounted to the robot arm and bed to monitor external vibrations, such as those initiating from the robot arm actuators, and to verify the effectiveness of the Sorbothane at preventing these external vibrations from being transferred to the implant and contaminating the signals at the implant sensors. Ultrasound transmission gel was used on each sensor to ensure a reliable sound transmission path by filling the space between the sensor diaphragms and the component each sensor was attached to.

The *in-vitro* test setup was designed to simply reproduce squeaking noises without incorporating bone or soft tissues into the apparatus. The exclusion of bone and soft tissues is an obvious limitation of the *in-vitro* test method since any damping or attenuation of vibrations induced by the implant will not be replicated. Therefore, it is important to note that any vibrations recorded during the *in-vitro* testing will likely have dominant frequency content that differs from that of the *in-vivo* AEs. Literature has shown frequencies of squeaking from similar *in-vitro* test setups to range between approximately 2.4-4.4 kHz and *in-vivo* squeaking to range between 0.5-2.5 kHz (Piriou et al., 2016). Numerical modelling has shown that the inclusion of bone and muscle around the implant components increases the stiffness of the system and

can account for the differences seen in vibration frequencies between *in-vitro* and *in-vivo* implant vibrations in the literature (Piriou et al., 2016). Consequently, it is not expected that the frequency content from any AEs detected *in-vitro* will exactly match that of AEs detected *in-vivo*. However, any squeaking AEs will be expected to have frequency content within the 0.5-4.4 kHz range. Furthermore, it is likely that the mechanism that induces squeaking is independent of the presence of bone and soft tissues since it is widely suggested by the literature that high friction between implant components is responsible for inducing squeaking (Currier et al., 2010; Hothan et al., 2011; Piriou et al., 2016).

3.5.3 Robot Arm Test Motions

The acetabular cup inclination was fixed at 45° to the horizontal to roughly approximate the inverse of the orientation of the pelvis, and the components were tested with two motion types; flexion-extension (FE) and rotation. Figure 3.7 shows a schematic of *in-vitro* setup with indicated FE and rotation ranges of motion. The FE angle was cycled sinusoidally between ±30° in the x-z plane from the vertical position shown in Figure 3.5. The cycling was performed at three different rates of approximately 1 Hz, 1.25 Hz, and 1.33 Hz to broadly replicate typical gait frequency. The rotation angle was cycled between ±80° from the +30° FE angle position, to simulate extreme movement, at a rate of approximately 1 Hz.

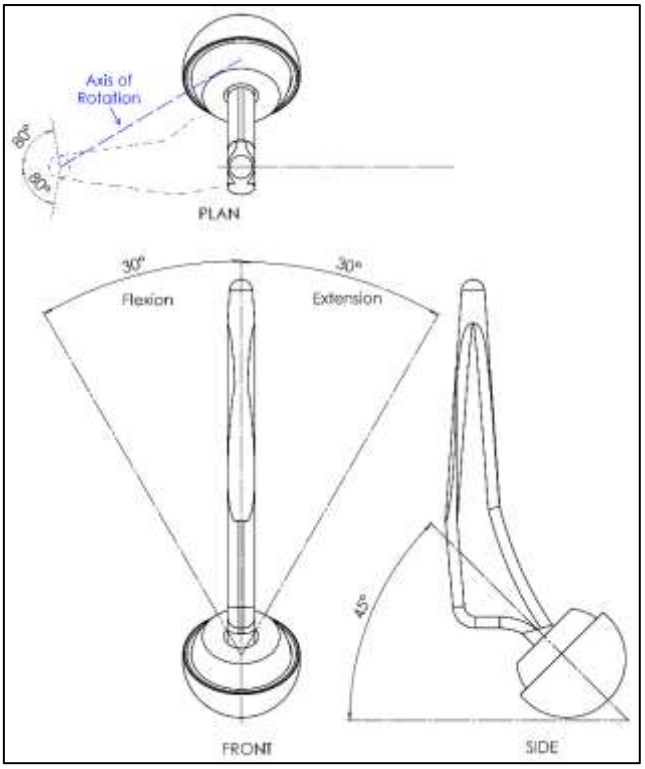

Figure 3.7: Schematic of in-vitro implant setup showing flexion-extension and rotation ranges of motion.

Prior to testing, the bearing surfaces of the femoral head and acetabular liner were wiped with acetone to remove any dust or oils present. Vertical offset of the robot end effector was used to ensure bearing surfaces were in good contact and a small load applied through the prosthesis. Due to the limitations in the robot, only low vertical loads up to 5kg were able to be applied to the implant. The low maximum implant loading provided by robot arm used was a limitation of this study but was unavoidable given the facilities available at the time of testing. Lubrication was not used during these tests as initial investigations found squeaking could not be reliably produced under lubricated conditions indicating the *in-vivo* occurrence of audible squeaking could be related to a loss of thin-film lubrication and increased friction as is suggested in the literature (Currier et al., 2010; Hothan et al., 2011; Piriou et al., 2016; Sariali et al., 2012; Walter et al., 2008; Weiss et al., 2012). In a lubricated implant in the human body, a high axial load may lead to the loss of the thin-film lubrication and increased friction. In the *in-vitro* testing, the complete absence of lubrication may offset the low loads applied by the robot arm. As with *in-vivo* testing, the sensor response from each motion (and different cycle rate) were

recorded separately to enable comparisons between the acoustic content for each specified motion.

3.5.4 Manual Implant Manipulation

The purpose of the manual manipulation was to investigate AEs generated from the trunnion Morse taper (TMT) interactions of the femoral components. Since the femoral head of the implants had only a press/impact-fit onto the cylindrical TMT of the femoral stem (see Section 2.2.3), it was hypothesised that the components could rotate relative to each other *in-vivo* if the femoral head had not been installed with sufficient impact force. Figure 3.8 shows a schematic of the femoral components and depicts the direction of rotation of the femoral head relative to the femoral stem.

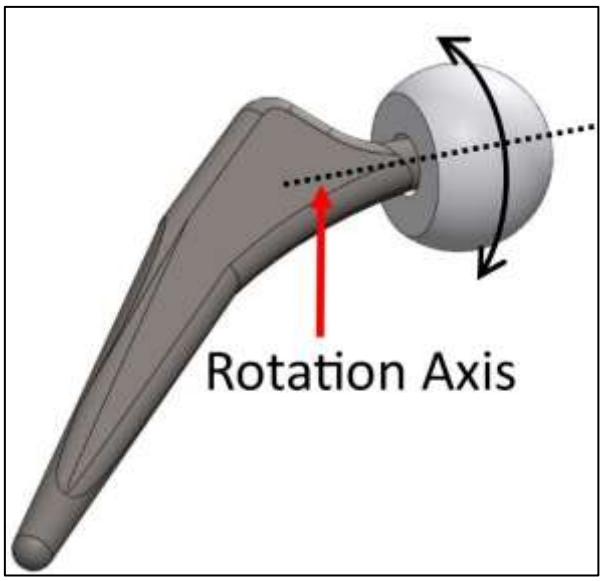

Figure 3.8: Schematic of femoral components indicating axis and direction of rotation.

To produce *in-vitro* AEs from the femoral TMT connection the femoral head was rotated back and forth by hand (in the direction depicted in Figure 3.8), while still connected to the femoral stem. To do this the femoral stem was held in one hand and the femoral head was rotated with the other hand. Both small and large rotation angles were tested with the small rotations being less than approximately ±5 degrees and the large rotations being up to ±45 degrees.

Additionally, both tight and loose fits of the femoral head were tested. Tight fits meant a large amount of friction was present between the Morse tapers of the components (simulating a femoral head installed with sufficient force) and hence a large amount of effort was required to rotate the femoral head. Loose fits meant the friction at the Morse tapers was greatly reduced and only little effort was required to rotate the femoral head thus simulating a femoral head installed with insufficient impact force. AEs were recorded during the rotations using a single sensor attached to the femoral stem in the same way that a sensor was attached to the femoral stem for the robot arm testing.

The rotation of the implant components was performed manually since an effective method to grip the femoral head and femoral stem with the robot arm to apply the desired rotations was not available within the timeframe of the testing. The manual manipulation introduces inconsistencies to the applied force, speed, and rotation of the implant components. Therefore, it is important to keep in mind that AEs recorded during the manual manipulations will be expected to have less consistency with respect to signal parameters when compared to AEs of the robot arm manipulations.

3.6 Summary

This chapter presented the details of the AE monitoring hardware and the data collection protocols that were developed for the *in-vivo* and *in-vitro* AE testing. The AE monitoring hardware utilises four ultrasonic sensors which, through the use of National Instruments data acquisition hardware, interface with LabVIEW software to record AEs on a laptop computer. *In-vivo* testing was conducted on consenting THR patients, with ethical approval from the New Zealand Upper South A regional ethics committee (URA/10/11/075). The *in-vivo* testing required study participants to perform basic motions (such as squatting, moving from standing

to sitting and *vice versa*, stair ascents and descents, walking, heel strikes, etc.) with the ultrasonic sensors attached to the skin surface from the iliac crest to the upper-femur. *In-vitro* testing was conducted on retrieved implants from study participants who had previously undergone *in-vivo* testing. The *in-vitro* testing utilised a 4-axis Epson RC20 robot arm to manipulate the retrieved implant components in motions that broadly replicated hip flexion-extension and rotation. Furthermore, *in-vitro* testing was conducted manually by rotating the femoral head about its rotational axis on the femoral stem (as shown in Figure 3.8) to investigate AEs from the interactions of the femoral components.

Chapter 4 General Analysis of Acoustic Emissions

4.1 Introduction

This chapter presents a summary of the demographics of the study participants and the initial results from the analysis of the recorded AE data. The primary focus of the chapter is to provide an introduction of the basic time and frequency domain appearances of the recorded AE signals from both the *in-vivo* and *in-vitro* testing. Figures showing visual representations of the signal waveforms are presented to assist in understanding the overall nature of the AE signal types. Since the AE technique has not been widely applied to the orthopaedic field, it is worth introducing the basic signal waveforms and associated frequency behaviour before attempting to analyse any diagnostic potential of the technique. However, this chapter will also explore some preliminary quantitative comparisons of the general frequency characteristics of *in-vivo* and *in-vitro* AEs. More specifically, comparisons of the general frequency behaviour of squeaking *in-vivo* and *in-vitro* AEs will be presented. This initial presentation of key characteristics and overall trends was deemed necessary before any more detailed analyses are undertaken in subsequent chapters. The clinical implications of the results of the comparisons will also be discussed. A subset of the results presented in this chapter have also been published in FitzPatrick et al. (2017b), Roffe et al. (2017) and FitzPatrick et al. (2017a).

4.2 Participant & Database Summary

The database of AE recordings continually increased as additional *in-vitro* testing was undertaken on implants and additional revision patients were recruited for *in-vivo* testing. At the time of writing, *in-vivo* data had been collected from 23 natural hips, for use as control AE data, and 117 total hip replacement (THR) patients. Of the hip replacement participants, 49 had

undergone subsequent revision surgery of their implants. Table 4.1 summarises the body mass index (BMI), participant age, and age of implants for the hip replacement participant cohort of the study. Table 4.2 provides the equivalent BMI, and participant age information for the control participant cohort.

Table 4.1: Summary of BMI, age, and implant age for hip replacement study participants.

	All			Male			Female		
	Average	St.dev.	Range	Average	St.dev.	Range	Average	St.dev.	Range
BMI (kg/m^2)	28.7	5.1	(18.5 - 52.8)	29.8	4.7	(21.8 - 52.8)	26.5	5.2	(18.5 - 35.4)
Age (yrs)	67.8	10.6	(23.9 - 91.4)	69.7	8.6	(48.6 - 91.4)	64.6	12.6	(23.9 - 82.7)
Implant Age (yrs)	10.6	6.1	(0.4 - 27.9)	10.9	6.3	(0.4 - 27.9)	10.0	5.5	(1.4 - 24.2)

Table 4.2: Summary of BMI and age for study control participants

	All			Male			Female		
	Average	St.dev.	Range	Average	St.dev.	Range	Average	St.dev.	Range
BMI (kg/m^2)	23.8	2.1	(20.2 - 27.3)	25.1	0.7	(24.5 - 26.0)	23.3	2.3	(20.2 - 27.3)
Age (yrs)	30.7	14.3	(21.0 - 58.6)	21.6	0.5	(21.0 - 22.1)	34.2	15.5	(21.2 - 58.6)

The purpose of the control participants was to provide AE data from persons with healthy natural hip joints with no history of hip joint trauma or injury in order to gain a reference for AEs that might occur from surrounding parts of the body. The summary of Table 4.2 shows the control participant cohort was made up of participants who were considerably younger than those participants of the hip replacement cohort. However, this age discrepancy of the control and hip replacement participant cohorts is not considered to impact the validity of AE comparisons between the cohorts as the purpose of the control AE data was met. The key underlying concept was to obtain a baseline for AEs produced by healthy natural hip joints. The two groups of control participants and hip replacement participants were not intended to be matched cohorts.

The hip replacement participants had five different bearing interface material combinations between them; ceramic-on-ceramic (CoC), metal-on-metal (MoM), metal-on-polyethylene (MoP), ceramic-on-polyethylene (CoP), and ceramic-on-metal (CoM). Furthermore, for 35 of the 117 THR participants the specific bearing surface material combination was unknown at the time of writing. The source of the implant bearing surface information was the participants' clinical notes and for the 35 unknown combinations the clinical notes had not been examined at the time of writing. Table 4.3 shows the number of artificial hips in each of the different bearing interface material combination categories for both the revised participants and those yet to undergo revision surgery.

Table 4.3: Number of artificial hips of each bearing interface material combination for both the participants who had undergone revision surgery and those that had not yet had revision surgery.

Bearing Interface	Number Revised	Number Unrevised	Total
CoC	11	30	41
MoM	5	1	6
MoP	23	0	23
CoP	7	1	8
CoM	1	3	4
Unknown	2	33	35
Total	**49**	**68**	**117**

The reason for revision was determined from the surgical notes of the 49 participants who had undergone revision surgery, and hence those participants were able to be grouped by revision reason. Table 4.4 presents the total number of artificial hips that had been revised for each of the revision reasons that were identified in the participant clinical notes. It should be noted that any participant can have multiple reasons for revision (such as having excessive wear, loosening, and pain) and therefore can be included in multiple categories. Hence, the sum of the categories in the Table 4.4 (72) is greater than the total number of participants/hips which had undergone revision surgery (49). Consequently, the possibility of participants having multiple reasons for revision needed to be considered in subsequent analysis.

Table 4.4: Number of artificial hips that were revised for each revision reason. Note participants can fall into multiple categories and hence the number of unique participants/hips is less than the total of all categories.

Revision Reason	Number of Hips
Noise	9
Wear	23
Loosening	21
Pain	12
Fractured Component	4
Dislocation	2
Infection	1
Sum of Categories	**72**
Total Number of Unique Revised Hips	**49**

The following analysis of AE data did not include the infection, dislocation, and fractured component categories due to the low number of hips in these categories. Furthermore, 11 of the 12 participants in the pain revision reason category were in at least one of the noise, wear, loosening, or fractured component categories. As a result, it was decided not to compare AE data in the pain category with the other categories as any trends in the pain category data were likely to be a combination of trends from the other categories.

4.3 In-vivo Acoustic Emissions

4.3.1 Overview

Initial examination of the *in-vivo* AE data involved the visual inspection of both time and frequency domain responses. Recall that the general definition of an AE event, as stated in Section 2.4.3, is an AE pulse that exceeds an arbitrary voltage threshold that is above the noise level and then decays back to the noise level. Furthermore, the AE event duration is generally defined as the time span from the first crossing of the arbitrary voltage threshold to the last crossing of the threshold before the signal amplitude returns to the base noise level. Upon visual inspection of the AE recordings it was found that the base noise level of the AE signal

recordings was very low, on the order of ±0.1 millivolts, relative to distinct AE pulses, some of which were on the order of ±2 Volts. Therefore, it was generally easy to visually identify that AE events had occurred without explicitly quantifying a voltage threshold. Chapter 5 presents a quantitative AE event signal analysis where an event detection algorithm was employed to identify distinct AE events and quantify their associated signal parameters across the entire AE dataset. However, the AE analysis presented in this chapter does not aim to quantitatively define the occurrence of specific AE events. Instead, this chapter presents observations from visually identified AE events that broadly represent the general behaviour of the AE signals. The primary focus of this chapter is to familiarise the reader with the basic time and frequency domain appearances of the recorded AE signals from both the *in-vivo* and *in-vitro* testing. In addition, a further focus of this chapter is to explore the broad frequency characteristics of the AE signals.

With that being said, it is useful to introduce two main initial observations at this point and verify them with a quantitative result from the analysis of Chapter 5. The first important observation from visual inspection of the time domain responses of the AE recordings was that the responses mostly contained short duration acoustic events, which were typically a few milliseconds (ms) in duration. However, it was observed that a small number of long duration acoustic events occurred, which were typically much greater than 20 ms in duration. To quantitatively verify the observation that the AE recordings contained mostly short duration (<20 ms) AE events, Figure 4.1a (from the analysis of Chapter 5) shows the empirical cumulative distribution functions (CDFs) of AE event durations of all identified *in-vivo* AE events for the THR participant and control groups.

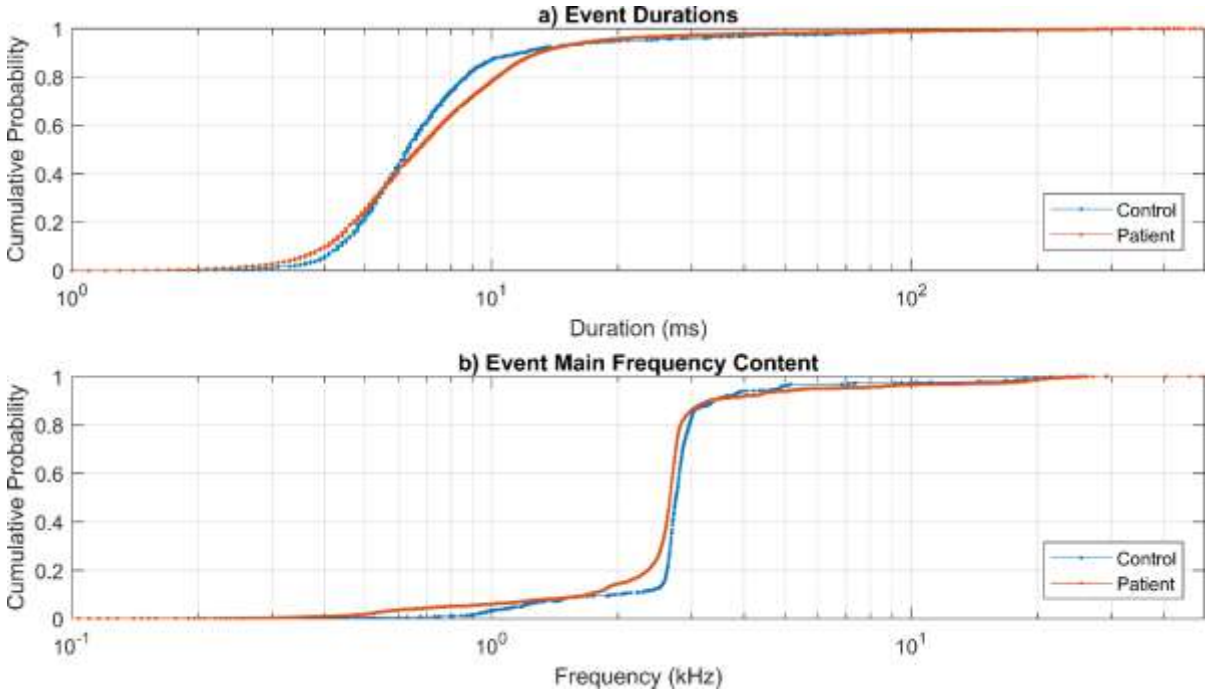

Figure 4.1: Empirical cumulative distribution functions (CDFs) of a) event durations and b) event main frequency content for the THR participant and control groups.

It was observed from the empirical CDF data of Figure 4.1a that approximately 95% of AE events identified in both the THR participant and control data sets had durations less than 20 ms. The long duration events, which made up only approximately 5% of observed acoustic events, were mainly considered to be audible squeaking events. Additionally, audible squeaking was only observed in participants with either ceramic-on-ceramic or metal-on-metal implant bearing surfaces which was consistent with literature (Imbuldeniya et al., 2013; Levy et al., 2015). Conversely, the short duration events were generally unnoticed by the naked ear during testing but were occasionally associated with 'clicking' or 'popping' sounds.

The second important initial observation was that the Fourier transform of the AE signals revealed broadly consistent main frequency content for the *in-vivo* data. For this initial analysis the Fourier transform for each signal was calculated using MATLAB's Fast Fourier Transform (FFT) algorithm applied to each AE signal separately. Recall from Section 2.4.3, the main frequency content corresponded to the frequency that had the greatest magnitude of the Fourier transform of an AE signal. Figure 4.1b (from the analysis of Chapter 5) shows the empirical

cumulative distribution functions (CDFs) of AE event main frequency content of all identified *in-vivo* AE events for the THR participant and control groups. It was observed that approximately 80% of all AE events identified in both the THR participant and control data sets had main frequency content between 1.7 and 3.4 kHz. The observed frequency range was in agreement with *in-vivo* fundamental frequencies of THR AEs found in similar studies. Walter et al. (2008) observed fundamental frequencies of *in-vivo* squeaking AEs in the range of 0.4-7.5 kHz for a cohort of 31 THR participants with a mean age and body mass index (BMI) of 54 years and 27 $kg \cdot m^{-2}$ respectively. Sariali et al. (2012) observed main frequency content between 1.45 kHz and 2.46 kHz from four participants with a mean age and BMI of 60.5 years and 22.4 $kg \cdot m^{-2}$ respectively. Currier et al. (2010) observed *in-vivo* squeaking of one male and one female participant to have fundamental frequencies of approximately 1.5 kHz and 2.5 kHz respectively.

As already stated, the results shown in Figure 4.1 were produced from the AE event analysis which is presented in Chapter 5. However, Figure 4.1 was included here to broadly introduce two general characteristics of the AE signals before the specific quantitative details of the event detection method are discussed in detail. The remainder of this chapter investigates the waveforms of some representative long and short duration AE event signals that were recorded from the THR study participants, control participants, and during the *in-vitro* testing. Additionally, the frequency characteristics of the AE signals in general (as opposed to the frequency characteristics of identified AE events) are explored and some clinical implications of these characteristics are discussed.

4.3.2 In-vivo Squeaking Events

The long duration events which were identified as squeaking usually had frequency responses showing a strong presence of harmonic frequencies. However, it was also observed that some

identified squeaking events had frequency responses that exhibited broad and noisy frequency peaks where harmonic frequencies were not so easily identified. Figure 4.2 shows the time and frequency domain magnitude responses of an *in-vivo* squeaking AE event from a select participant (Participant 1). It was observed that the event duration was approximately 110 ms and the overall magnitude trend of the waveform over time was symmetric: increasing for the first half of the duration and decreasing for the second. Additionally, the frequency response was observed to have a very large relative magnitude and narrow bandwidth peak at approximately 1 kHz with three additional lower magnitude peaks at the associated harmonic frequencies of approximately 2, 3, and 4 kHz.

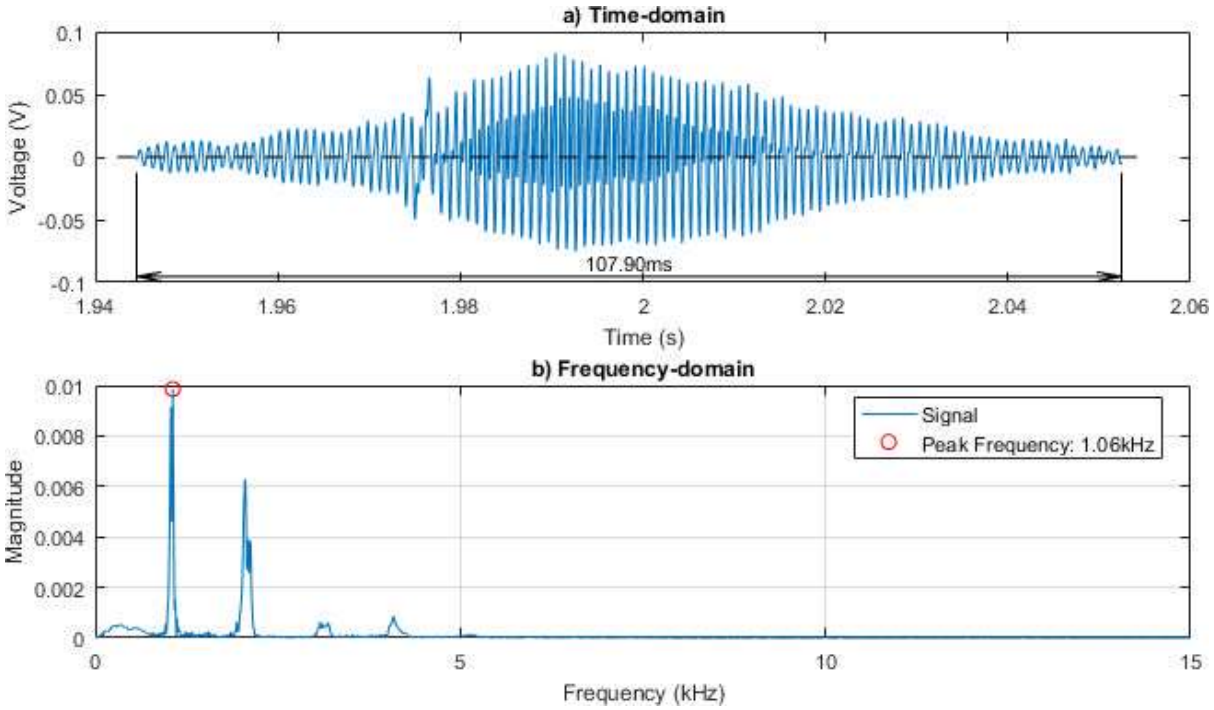

Figure 4.2: (a) Time-domain and (b) frequency-domain response plots of an in-vivo AE event from Participant 1, recording 6.

Conversely, Figure 4.3 shows the time and frequency domain magnitude responses of a second squeaking event from the same select participant as that of Figure 4.2. It was observed that the event duration was approximately 570 ms which was over five times longer than that of the event of Figure 4.2. The overall magnitude trend of the waveform was observed to be much less consistent than that of Figure 4.2 but the peak magnitude occurred near the centre of the

event duration as it did in Figure 4.2. The frequency response was observed to have a very large relative magnitude and narrow bandwidth peak at approximately 1 kHz, similar to Figure 4.2. The same presence of lower magnitude peaks at the associated harmonic frequencies of approximately 2, 3, and 4 kHz was also observed in Figure 4.3 but the harmonics are not as clear as they were in Figure 4.2. An additional peak at the fourth harmonic frequency of approximately 5 kHz was also observed.

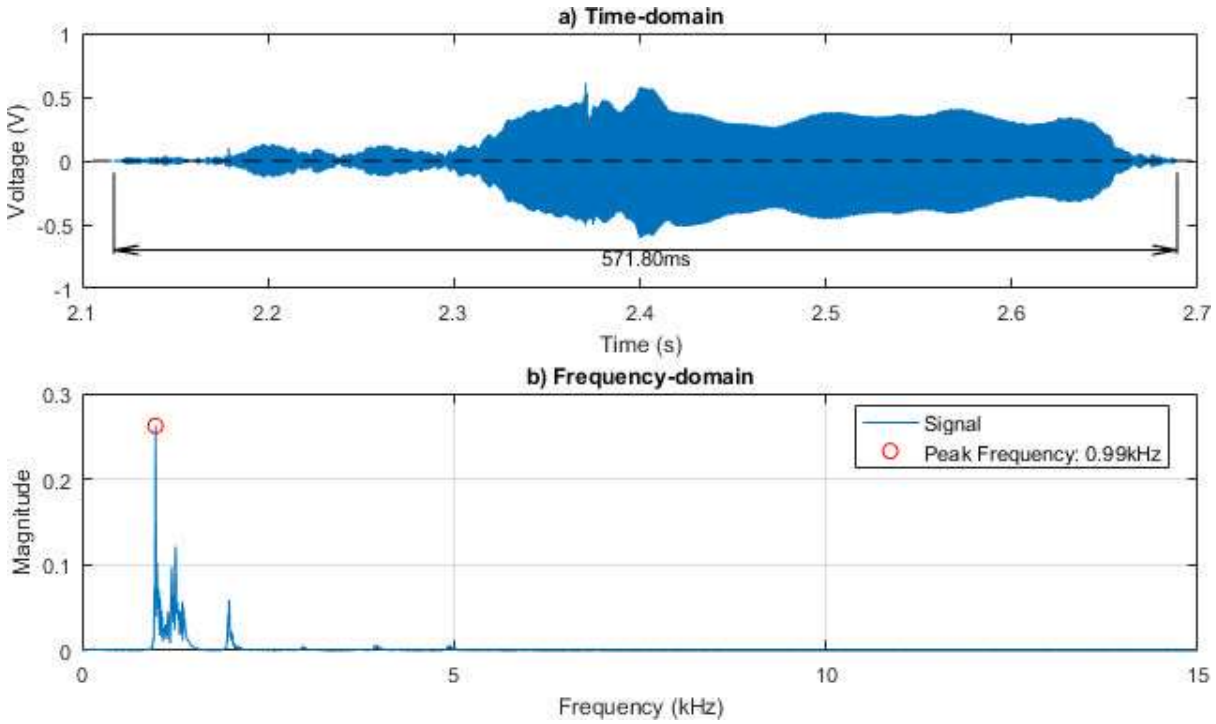

Figure 4.3: (a) Time-domain and (b) frequency-domain response plots of a second in-vivo AE event from Participant 1, recording 6.

The AEs shown in Figures 4.2 and 4.3 show broadly similar behaviours. However, these two AEs do have a number of marked differences. In general, the voltage amplitude of the AE event of Figure 4.3 is approximately 10 times larger than that of the AE event of Figure 4.2. As mentioned previously, the AE event of Figure 4.3 is approximately five times longer in duration than that of Figure 4.2. Furthermore, the time domain response of Figure 4.3 indicates less consistent voltage amplitudes over time compared to that of Figure 4.2. The frequency domain response of Figure 4.3 has its main frequency peak at approximately 1 kHz and an additional and broad magnitude peak centred at approximately 1.25 kHz which was not present in Figure

4.2. This additional frequency magnitude peak has the second greatest magnitude of the frequency peaks of Figure 4.3 and is clearly not a harmonic peak. Hence, the presence of this additional peak indicates a secondary vibration was simultaneously present during the recording presented in Figure 4.3.

Figure 4.4 shows the time and frequency domain responses of an *in-vivo* squeaking AE event from a different participant (Participant 27). It was observed that the main event duration of Figure 4.4 was approximately 190 ms and the signal waveform contained distinct pulses towards the end of the event which were not present at the beginning. These distinct pulses observed from the signal may indicate a stick-slip type friction motion at the implant main bearing interface. The frequency response was observed to have three very broad and noisy magnitude peaks at approximately 1, 2, and 3 kHz with the greatest magnitude occurring at the 2 kHz peak. Given the approximate 1 kHz spacing of the frequency peaks, it is likely that the event of Figure 4.4 contains harmonic frequencies. However, since the frequency response was observed to have relatively noisy magnitude peaks it was possible the signal contained superposition or interference of AE signals. Figure 4.4b shows much less clearly defined peak frequencies and a broader range of frequency content than those seen in Figures 4.2 and 4.3.

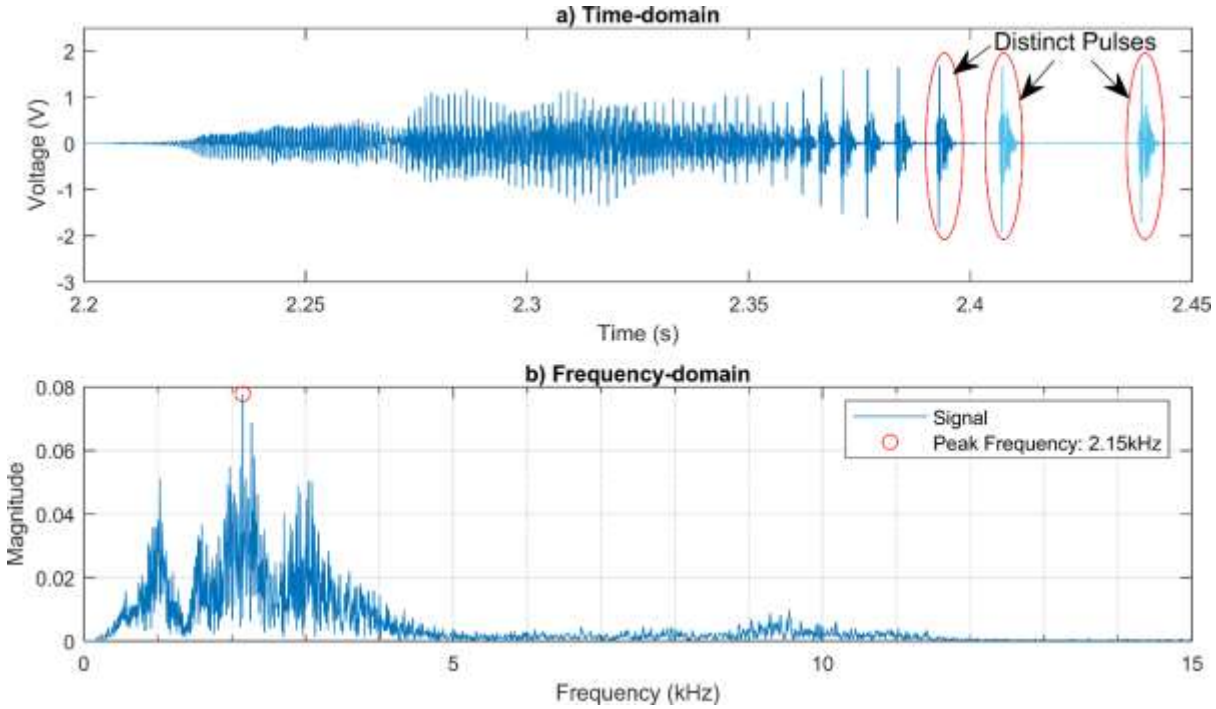

Figure 4.4: (a) Time-domain and (b) frequency-domain response plots of an in-vivo AE event from Participant 27, recording 1.

Superposition of AE signals or changes in the properties of the generated acoustic waveform could result from AEs emanating simultaneously from different sources, reflections and refractions of AEs from a single or multiple sources, or changes in the properties of an AE source over time. Further analysis of the signal was necessary to investigate the possibility of signal superposition of multiple AE signals or temporal changes in observed acoustic waveform properties. One method for further analysis of the signal was to investigate the frequency behaviour of the signal over time by performing short time Fourier transforms (STFTs).

The STFT of a signal splits the AE data recording into short time windows (of 0.01 seconds in this case) and shows the frequency response for each of these time windows on one plot. The STFT was found by applying MATLAB's FFT algorithm to each short time window and plotting the resulting FFT against time yielding a three-dimensional plot known as a spectrogram. Each individual time window had a 10% duration overlap with the previous window and a Hanning windowing function was applied to each window prior to performing

the FFT. The magnitude of each frequency at each window in time is represented by the colour of the plot and is expressed in decibels (dB) given by the corresponding colour bar. Figure 4.5a and 4.5b show the time domain response and the spectrogram respectively of the same AE event as shown in Figure 4.4.

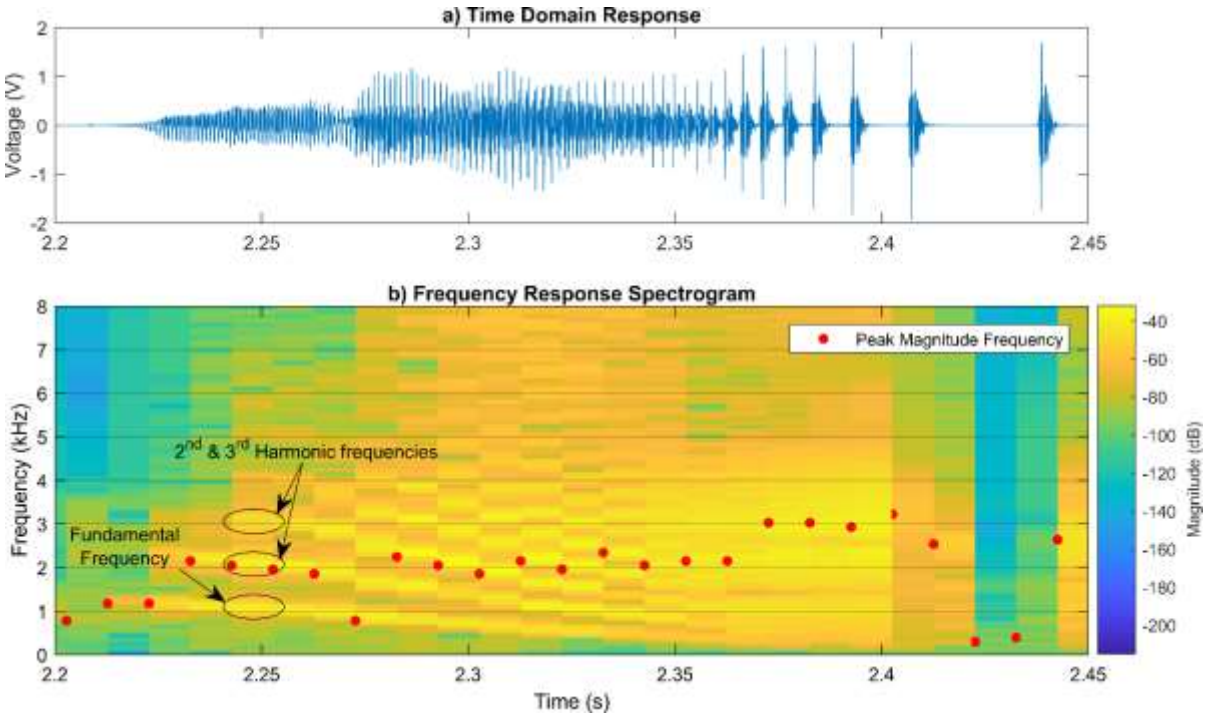

Figure 4.5: a) Time-domain response and b) spectrogram plots of same in-vivo AE recording from Figure 4.4.

From Figure 4.5b it was observed that the magnitude of the frequency response of the AE signal changed over time. It can be observed from Figure 4.5b that the main frequency content of the signal (i.e. largest magnitude peak within that time window), indicated by the red dots, ranges between approximately 0.8 kHz at the beginning of the signal and approximately 3.2 kHz towards the end of the signal. Furthermore, at the beginning of the signal the main frequency content was observed to fluctuate around approximately 1 kHz. The main frequency content then shifted to fluctuate between approximately 1.8-2.2 kHz for the majority of the signal duration. Finally, towards the end of the signal, the main frequency content shifted again to fluctuate around approximately 3 kHz. Thus, these main frequency content observations from the spectrogram are, to some extent, consistent with the average frequency response of the

whole signal, which was shown in Figure 4.4b. From Figure 4.4b three broad frequency magnitude peaks were observed to be centred on approximately 1, 2, and 3 kHz, with the 2 kHz peak being considered the main frequency content of the signal because it had the largest magnitude of the three peaks. The spectrogram confirmed the observations of Figure 4.4b but additionally was able to identify that the main frequency content of the signal changed over time and that all three of the identified peaks corresponded to the main frequency content at separate windows in time. This example shows that a single FFT of an entire event may mask the varying behaviour throughout an event. Conversely, the spectrogram provides additional information about shifting frequencies throughout a recording as implant load, articulation angles, and other variables change as the participant moves.

Another observation from Figure 4.5b was the presence of harmonic frequencies. The harmonic frequencies were indicated by increased frequency magnitude values (bright yellow regions) occurring at frequency multiples in a specific time window. For example, in Figure 4.5b at the time window that included 2.25 seconds, an increased frequency magnitude was observed that occurred at approximately 1 kHz, indicating the fundamental or 1^{st} harmonic frequency, with additional bright yellow regions at the multiples of 2 and 3 kHz, indicating the 2^{nd} and 3^{rd} harmonic frequencies respectively (see annotations in Figure 4.5b). It was observed from Figure 4.5b that, for some of the time windows, up to five or more harmonic frequencies were present in the signal. Furthermore, it was evident from the spectrogram of Figure 4.5b that the fundamental frequency began at approximately 1.2 kHz at the start of the signal and then decreased in frequency approximately linearly over time until it was around 0.4 kHz towards the end of the signal. Consequently, the higher harmonic frequencies also exhibited decreases in frequency over time. This observation of decreasing harmonic frequencies over time is important because it helps to explain the broad frequency magnitude peaks that were seen in the average frequency response of the whole signal in Figure 4.4b.

The average frequency response of the signal (Figure 4.4b) showed noticeable frequency magnitude content from approximately 0.5 kHz through to almost 5.0 kHz. From observing the bright yellow regions of the spectrogram of Figure 4.5b it is evident that the spectrogram does support the observation of noticeable frequency content across the range of 0.5-5.0 kHz, but not for every moment in time. In addition, as identified previously, the spectrogram makes it evident that over the duration of the AE the main frequency content fluctuated around either 1, 2, or 3 kHz with 2 kHz being the main frequency for the highest proportion of the duration. Hence, it follows that the average frequency response showed a general increase of frequency content in the 0.5-5.0 kHz range, greater frequency magnitudes around 1, 2, and 3 kHz, and the greatest frequency magnitude occurring at 2 kHz.

In general, similar behaviours of harmonic frequencies changing over time were also seen when observing the spectrograms of other audible squeaking AEs. One important factor reported in literature that has been shown to cause changes in the frequency behaviour of squeaking AEs was implant loading. Hothan et al. (2011) demonstrated from *in-vitro* testing that increasing the axial load on THR implants between 200 Newton (N) and 1800 N caused an increase in the squeaking frequency from approximately 3.1 kHz to approximately 4.4 kHz. Given that the *in-vivo* squeaking AEs occurred during dynamic participant motions, it is likely that the THR implants undergo significant changes in loading as a participant moves during squeaking events. The load on a hip joint during normal gait is in the approximate range of 0-400% body weight (Brinckmann et al., 2002). However, while the general motion of the participant was known (i.e. walking, stand from sit, stair ascent, etc.) for each squeaking event, the specific stage of the motion was not known (e.g. percentage of gait cycle for walking tests). Hence, for AEs collected using the protocol of Section 3.4, it was not possible to accurately determine a loading profile of the THR implants during squeaking AEs to verify any influence of implant loading on the observed squeaking frequencies. Nevertheless, the additional insight that the

spectrogram provided concerning the temporal frequency behaviour of the AE squeaking signals was useful as it demonstrated that average frequency responses could be misleading. Therefore, metrics such as main frequency content that have been acquired from the average frequency response of a squeaking AE should be used with caution in any further analyses.

4.3.3 Method for Determining Squeak Fundamental Frequency

The previous section identified that squeaking acoustic emission events generally exhibited harmonic frequencies. It was also observed that the main frequency content of the squeaking AEs often fluctuated between different harmonic frequencies over the duration of a squeak. These fluctuations in main frequency content result in average main frequency content values that can misrepresent the frequency behaviour of an AE event. Therefore, for squeaking AE events it would be more advantageous to find and compare fundamental frequency content when harmonic frequencies are present. One technique to identify the fundamental frequency of squeaking AE events was to employ a cepstrum. This section briefly introduces the concept of a cepstrum and presents results of the cepstrum technique applied to squeaking AEs. For a more detailed analysis of the cepstrum technique see the published conference proceeding covering this work (FitzPatrick et al., 2017a) in Appendix A.

The cepstrum method is a variation on traditional frequency domain analyses, with some advantages when multiple strong harmonics are present. Bogert et al. (1963) defined the power cepstrum of a signal as “the power spectrum of the logarithm of the power spectrum” (Randall, 2013). The concept of a cepstrum was developed while trying to identify echoes in seismic signals. Oppenheim and Schafer (2004) showed that a signal containing an echo can be represented as the summation of the original signal with a scaled and time delayed version of the original signal. The Fourier transform of the signal containing the echo is then equal to the

Fourier transform of the original signal multiplied by a periodic function that is the spectrum contribution of the echo (Oppenheim & Schafer, 2004). Then, by the rules of logarithms, the logarithm of the Fourier transform becomes the summation of the log spectrum of the original signal and a periodic component which has a 'fundamental frequency' of the echo delay (Oppenheim & Schafer, 2004). See Appendix A for the associated equations.

Given that periodic components appear as sharp peaks in the spectrum of a signal, if the Fourier transform of the aforementioned logarithm is computed it will contain a peak at the position corresponding to the time that the echo occurred in the original signal (Oppenheim & Schafer, 2004). However, the resulting spectrum, although a Fourier transform of a signal, is not in the frequency domain, as it has a unit of time as the independent variable, but it is also not in the time domain as the signal is different to the original waveform. Thus, Bogert et al. (1963) named this domain the *quefrency* domain and the spectrum of logarithm as the cepstrum since they found they were "operating on the frequency side in ways customary on the time side and vice versa" (Oppenheim & Schafer, 2004).

While finding the presence of echoes is useful in some applications it is not immediately useful for AE analysis of THR implants. However, cepstrum analysis has another property which is of more immediate use for analysing THR AEs. Cepstrum analysis has the ability of identifying the fundamental frequency of a vibration when multiple harmonics are present. Consider a sinusoid with decayed first and second harmonics of the following form:

$$x(t) = \beta_1 \sin(2\pi f_1 t) + \beta_2 \sin(2\pi f_2 t) + \beta_3 \sin(2\pi f_3 t) \tag{4.1}$$

where $\beta_1 = 1$, $\beta_2 = 0.3$, $\beta_3 = 0.1$, $f_1 = 2Hz$, $f_2 = 4Hz$, and $f_3 = 6Hz$. Figure 4.6a shows the log-magnitude Fourier power spectrum of a Hanning windowed version of the signal defined in Equation (4.1). As expected, it can be observed from Figure 4.6a that three frequency peaks

are present at the appropriate frequencies of 2, 4, and 6 Hz. By taking the inverse Fourier transform of the log-magnitude signal in Figure 4.6a the cepstrum in Figure 4.6b is produced. However, the signal in Figure 4.6b does not immediately appear useful and so Figure 4.6c was created by replotting the data of Figure 4.6b using inverse quefrency values. By using the inverse quefrency values the horizontal axis of Figure 4.6c has units of Hz and it can be observed that a sharp peak is present at 2 Hz. Hence, the cepstrum has successfully identified the fundamental frequency present in Equation (4.1) of 2 Hz, despite the strong presence of harmonics.

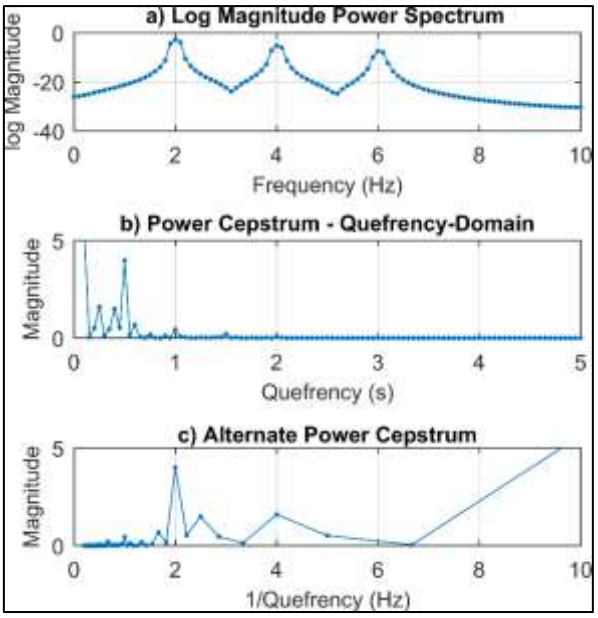

Figure 4.6: a) Log-magnitude Fourier power spectrum of a Hanning windowed sinusoid containing two harmonic frequencies, b) cepstrum of the sinusoid, and c) alternative representation of the cepstrum to show units of Hz on the horizontal axis.

The ability to identify the fundamental frequencies of a signal containing harmonics is a useful property of the cepstrum analysis. As mentioned in Section 4.3.2, some squeaking AE signals from THR participants show a strong presence of harmonic frequencies. Therefore, the cepstrum analysis can be applied to assist in the comparison and correlation of squeaking AE signals of THR participants. Additionally, throughout the duration of audible squeaking, the position of the fundamental frequency tended to show a gradual shift of up to 1 kHz in some

cases. As a result, a short-time (ST) cepstrum analysis was applied to better reflect the transient frequency content of the AEs.

The existing THR AE signals have been sampled at a rate of 100 kHz giving a signal time-step of 0.01ms. To perform the ST cepstrum analysis MATLAB software was used to apply an algorithm to individual THR AE signals to generate corresponding ST cepstrum signals. The algorithm to compute a ST cepstrum signal was as follows:

1. Divide the signal array into 10ms windows of time and treat each time window as a separate signal between 0 and 10ms (That is, each signal window's time axis array will be from 0 to 10ms in steps of 0.01ms.).
2. Perform a Fast Fourier Transform (FFT) on the first time-domain window signal and compute the modulus of the resulting signal. Note that this new signal array is the frequency-domain response of the time-domain signal and has a corresponding frequency value array from 0 to 100 kHz in steps of 100 Hz ($dF = 100kHz/1000samples = 100Hz$).
3. Compute the natural logarithm of FFT result from Step 2.
4. Perform an inverse FFT of the result from Step 3.
5. Compute the modulus of the result from Step 4. This new signal array is the quefrency-domain response and because a FFT and consequent inverse FFT have been performed, the corresponding quefrency value array is the same as the time-domain time-step values from 0 to 10ms with a quefrency-step of 0.01ms.
6. Repeat Step 2 to 5 for each of the remaining time-domain signal windows.
7. Calculate the array of equivalent frequency values by taking the inverse of the quefrency value array (i.e. 1/*quefrency*).

To present the results, a 3-dimensional ST frequency-domain plot (known as a spectrogram) and 3-dimensional ST cepstrum plot were created for each AE signal. The spectrogram was created by plotting each frequency magnitude result array (Step 2 of the algorithm) on the z (out-of-page) axis against the corresponding frequency values on the y (vertical) axis and at the appropriate window position in time on the x (horizontal) axis. The window positions on the x-axis corresponded to the time at the centre of each window from the original time-domain AE signal (for example, the result arrays for the first, second, and third signal windows were placed at 5, 15, and 25ms on the x-axis respectively). The ST cepstrum was created by plotting each quefrency magnitude result array (Step 5 of the algorithm) on the z-axis against the corresponding equivalent frequency values (Step 7 of the algorithm) on the y-axis and at the appropriate window position in time on the x-axis. The resulting 3D plots are shown here as 2D colormaps.

Figure 4.7a shows the time-domain response of an audible squeak from a THR participant. Figure 4.7b shows the ST frequency-domain response (spectrogram) of the AE signal from Figure 4.7a with lighter colours representing higher magnitude of frequency content and red dot markers indicating the largest magnitude frequency content. The spectrogram shows the frequency content of the time-domain signal over time and hence any transient frequency behaviour can be identified. For the purpose of this cepstrum analysis the absolute magnitudes of the frequency responses (and quefrency responses) are not important and therefore the absolute magnitude scale (colorbar) has been included only to indicate the relative magnitude of one colour to the next; this analysis is primarily concerned with the locations of peak frequency content and fundamental frequencies which can be found by considering relative magnitudes.

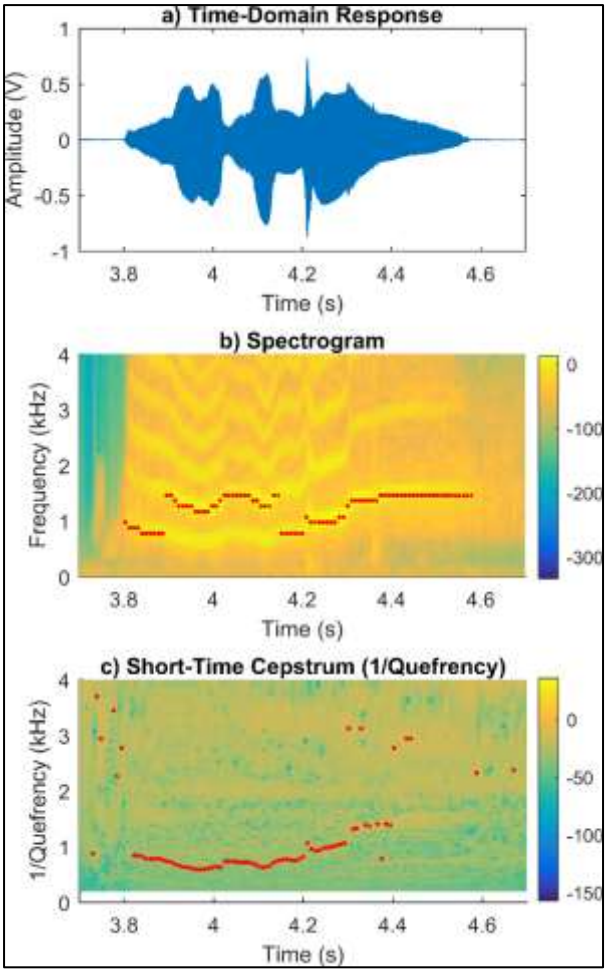

Figure 4.7: a) Time-domain response of an AE audible squeak from a THR patient; b) corresponding Frequency-domain and c) quefrency-domain responses showing lighter colours representing higher magnitude and peak magnitudes denoted by red dot markers.

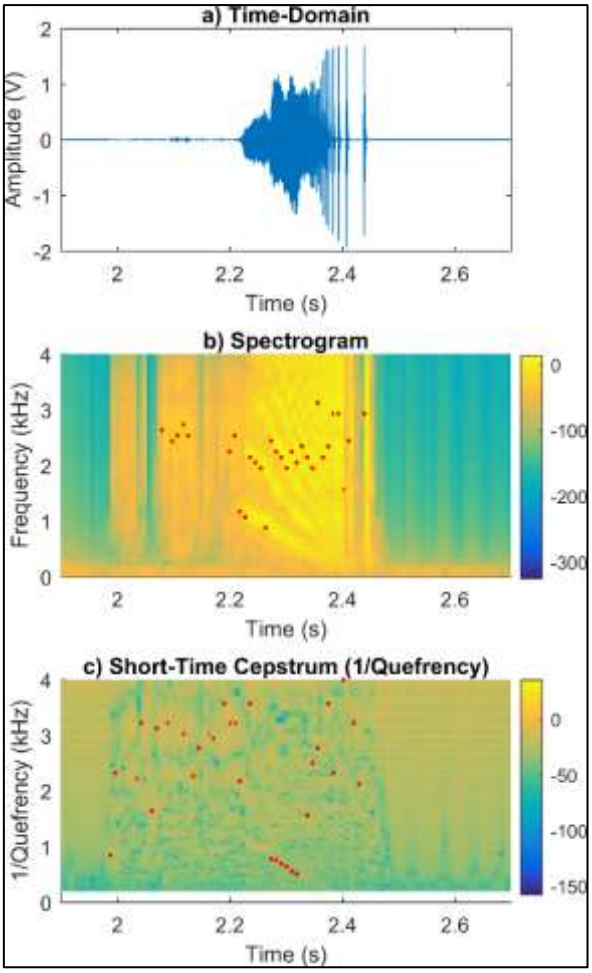

Figure 4.8: (a) Time-domain plot of an audible AE squeak from a second THR patient; (b) corresponding frequency-domain and (c) quefrency-domain responses. Lighter colours represent larger magnitude and red dots denote peak magnitudes.

It was observed from Figure 4.7b that there was a strong presence of harmonics in the signal and the fundamental frequency changes substantially over time (having a range of approximately 1 kHz from approximately 0.7 to 1.7 kHz). The frequency of largest Fourier magnitude (indicated by red dot markers) in Figure 4.7b was observed to move from the fundamental frequency to the first harmonic for a short time and then return to the fundamental frequency. This shifting of the largest magnitude frequency between the fundamental and associated harmonics was observed in a number of the THR AEs when audible squeaking was present. This shifting between harmonics creates difficulties in the identification of both the

frequency of the fundamental vibration mode and the associated variation in this frequency over time.

Figure 4.7c shows the ST cepstrum response of the same AE of Figure 4.7a and 4.7b, again with lighter colours representing larger magnitudes. Figure 4.7c shows that the peak magnitudes, indicated by red dot markers, compared very well with the fundamental frequency peak of the corresponding spectrogram of Figure 4.7c during the period when strong harmonics were present. During the time windows where the harmonics were less pronounced or non-existent the cepstrum peak magnitudes appeared to demonstrate much higher variability.

Figure 4.8a shows the time-domain response for an AE audible squeak from a second THR patient and Figure 4.8b shows the corresponding spectrogram. The spectrogram shows that the majority of the peak frequency content (indicated by the red dot markers) was scattered between 2 to 3 kHz. However, between 2.2 and 2.4 seconds the spectrogram shows the fundamental frequency linearly decreases from approximately 1.2 kHz to approximately 0.2 kHz with the presence of a number of harmonics. Therefore the frequency of largest Fourier magnitude in Figure 4.8b alone does not necessarily give an accurate representation of the frequency behaviour of the AE, as illustrated by the scattered red dot markers.

Figure 4.8c shows the ST cepstrum response of the AE from Figure 4.8a. It was observed that for regions of the signals where harmonics were not strong or well-defined, the peak cepstrum magnitudes appeared scattered across the quefrency axis. However, in the region between approximately 2.25 and 2.35 seconds the peak cepstrum magnitudes showed a linear decrease from approximately 1 kHz to approximately 0.5 kHz. Figure 4.9a shows the time-domain response for an AE audible squeak from a third THR patient and Figure 4.9b shows the corresponding spectrogram. It was observed from the spectrogram that the largest magnitude

frequency between approximately 0.7 and 0.9 seconds was approximately 3 kHz but showed a small increasing trend with time. The main frequency content from 0.9 seconds dropped to approximately 1.7 kHz and was steady for the remainder of the squeak. Figure 4.9c shows the ST cepstrum response of the AE from Figure 4.9a. It was observed that between approximately 0.8 and 0.9 seconds the peak cepstrum magnitudes were approximately 1.5 kHz and showed a similar shallow increasing trend to that of the peak spectrogram frequencies during the same time period.

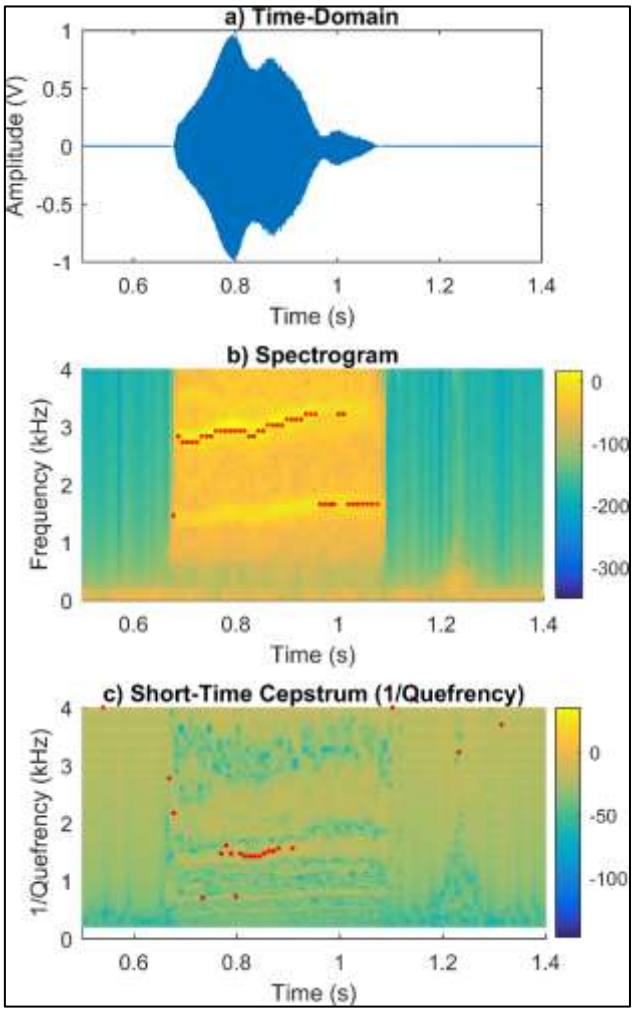

Figure 4.9: a) Time-domain plot of an audible squeak from a third THR patient, (b) corresponding frequency-domain and (c) quefrency-domain plots. Red markers are peak magnitudes.

The spectrogram results of the three different THR patient's audible squeak signals (Figure 4.7b, Figure 4.8b, and Figure 4.9b) had peak frequency content that varied substantially. Some of the variation in peak frequency content was observed to be due to peak magnitudes switching

between the fundamental frequency mode and various harmonics over time. This behaviour can be explained by considering that a larger amount of energy would be required to excite a higher vibration mode of a component and hence if different amounts of vibrational energy are present at different stages of the squeak different vibration modes would be excited more strongly. To support this idea, Figure 4.9a and Figure 4.9b show that the period of time after the peak frequency content drops from the 1st harmonic (which was approximately 3 kHz) to the fundamental frequency (which was approximately 1.5 kHz), the time-domain amplitude is significantly reduced and hence less vibrational energy is present.

During the time periods when the peak frequency content of the AE signals was consistently identifying either the fundamental or a given harmonic, some variation in the peak frequency was still observed. This variation was generally limited to a maximum of 200 Hz between time windows. One possible explanation for this additional variation is the influence of implant positioning, muscle forces and associated implant loads during the patient gait cycle. The squeaking signals were recorded while the patient was performing basic orthopaedic test motions such as walking, standing and sitting, and stair ascents. Therefore it is very likely that implant joint angles and loadings changed significantly during the squeak. Literature concerned with squeaking of THRs has found that squeaking frequencies tend to increase with increased implant loads (Hothan et al., 2011). An increase in implant loadings may affect factors such as the stiffness of the entire implant-bone system and consequently modify the natural vibration frequencies of the system.

The results from the ST cepstrums showed that when strong harmonics were present in the AE signals, the peak magnitudes of the cepstrums were closely aligned with the fundamental frequencies seen in the spectrogram results. During these periods where strong harmonics are present, the cepstrum analysis performed well at collapsing the higher harmonics of the signal

and identifying the fundamental vibration frequency. The identification of the fundamental vibration frequency allows for more reliable comparisons of squeaking frequencies between separate AE signals. Prior comparisons of AE signals compared the peak magnitude frequency content which exhibited substantial variations due to the harmonics present in the signal. Comparisons between AE signals using the frequencies identified by the cepstrum analysis has the potential to more reliably categorise AE signals based on the fundamental mode and thus assist in identifying signals that characterise particular implant failure modes or conditions.

One of the main limitations of the cepstrum is its inability to isolate frequencies of significance when strong harmonics are not present. This limitation is easily observed in the scattered peak magnitude markers of Figure 4.7c, Figure 4.8c, and Figure 4.9c for periods of time where the corresponding spectrograms show weak harmonic frequencies. It might be expected that the cepstrum would be able to identify the peak magnitude frequency content during time periods without harmonics; however, this was not the case and the associated frequencies of the peak cepstrum magnitudes during these periods had no obvious significance. It is therefore important to carefully consider the frequency content of any signal before applying a cepstrum analysis in order to interpret the results accordingly.

Another limitation of the cepstrum is that while the quefrency resolution is constant, the equivalent frequency resolution is not constant. The equivalent frequencies of the cepstrum are related to the quefrencies through an inverse relationship (i.e. frequency = 1/quefrency) and hence the frequency resolution of the cepstrum is very high at low frequency and decreases as frequency increases. This trend means that the cepstrum may produce distorted result if the frequencies of interest have a large bandwidth. In the case of the audible squeaks of Figure 4.7, Figure 4.8, and Figure 4.9 a large bandwidth is not a significant issue as the frequencies of interest are below 4 kHz. The coarsest frequency resolution up to 4 kHz is approximately

150 Hz which is sufficient for useful frequency comparisons between squeaks. However, the equivalent frequency resolution must be considered if the frequencies of interest become higher than 4 kHz, or if the length of the cepstrum time window decreases, as this will also affect the size of the quefrency resolution.

A further limitation of the cepstrum is its sensitivity to signal noise. The cepstrum is defined as the "spectrum of the log-magnitude spectrum of a signal" (Randall, 2013). The spectrum of a signal will inevitably contain some level of noise. When the log-magnitude spectrum is computed and then treated as a linear scale in decibels, any noise will be amplified along with the signal. When the inverse Fourier transform is applied to the log-magnitude spectrum the amplified noise will much more significantly affect the result. Higher sampling rates will also have a similar effect as they will include high frequency noise in the signal, which may have otherwise not been recorded. Much of the scatter observed in Figure 4.7, Figure 4.8, and Figure 4.9 can be attributed to time windows without major AE content, resulting in primarily ambient noise with a low signal-to-noise ratio.

4.3.4 In-vivo Non-squeaking Events

Worldwide research interest in the occurrence of THR implant squeaking has meant that special attention has been given to the study of THR implant squeaking. The origin of squeaking is still debated among researchers and is repeatedly reported as multifactorial. Implant squeaking produces strong AE signals which are easy to record and to distinguish from other recorded AE signals. Therefore, the analysis of THR implant squeaking is a good starting point for investigation of implant AEs in general. However, it should be noted that audible squeaking is not considered a major problem of THR implants. As identified in Section 2.2.4, a recent meta-analysis found the incidence of squeaking in ceramic-on-ceramic THRs to be 4.2%, of which

the incidence of subsequent revision for squeaking was 0.2% (Owen, D. H. et al., 2014). Implant failures such as loosening and repeated dislocations have a much greater prevalence in THR participants and squeaking is not usually associated with these failure mechanisms. Therefore, this study is not focussed exclusively on THR implant squeaking as is demonstrated in the following sections.

The large majority (approximately 95%) of the *in-vivo* acoustic events observed from initial observations of the time domain signals had durations below 20 ms. In contrast to audible squeaking events, most of these short duration (<20 ms) events were not heard by the naked ear during the *in-vivo* testing. A very small number of audible 'clicks' or 'pops' did occur during testing but it was not practical to keep track of which specific acoustic event matched an audible sound during testing (in some cases, a note was taken of the specific test recording during which an audible sound occurred, but no particular acoustic signal event or fluctuation could be reliably associated with the audible noise). Nonetheless, being able to confirm the presence of audible short duration acoustic events in a particular acoustic recording was of less importance to the study than noting the presence of audible squeaking.

Figure 4.10 shows the time and frequency domain responses of an *in-vivo* short duration event from a select participant (Participant 1). It was observed that, at 6.7 ms, the event duration was significantly lower than those of the squeaking events in Section 4.3.2 (Figures 4.2-4.4). The overall magnitude trend was observed to be similar to the behaviour that would be expected from an impulse excitation, i.e. peak magnitude is reached relatively quickly (short rise time) and then the oscillation amplitude decays gradually to background noise levels. An important observation to note was that the peak amplitude of this acoustic event was approximately 0.15 V which was comparable to the amplitude of the first squeak of Section 4.3.2 in Figure 4.2. The frequency response of the event of Figure 4.10 was observed to have a main peak at

approximately 3 kHz and did appear to contain shallow broad harmonic frequency peaks at approximately 6 and 9 kHz, potentially indicating the presence of harmonics.

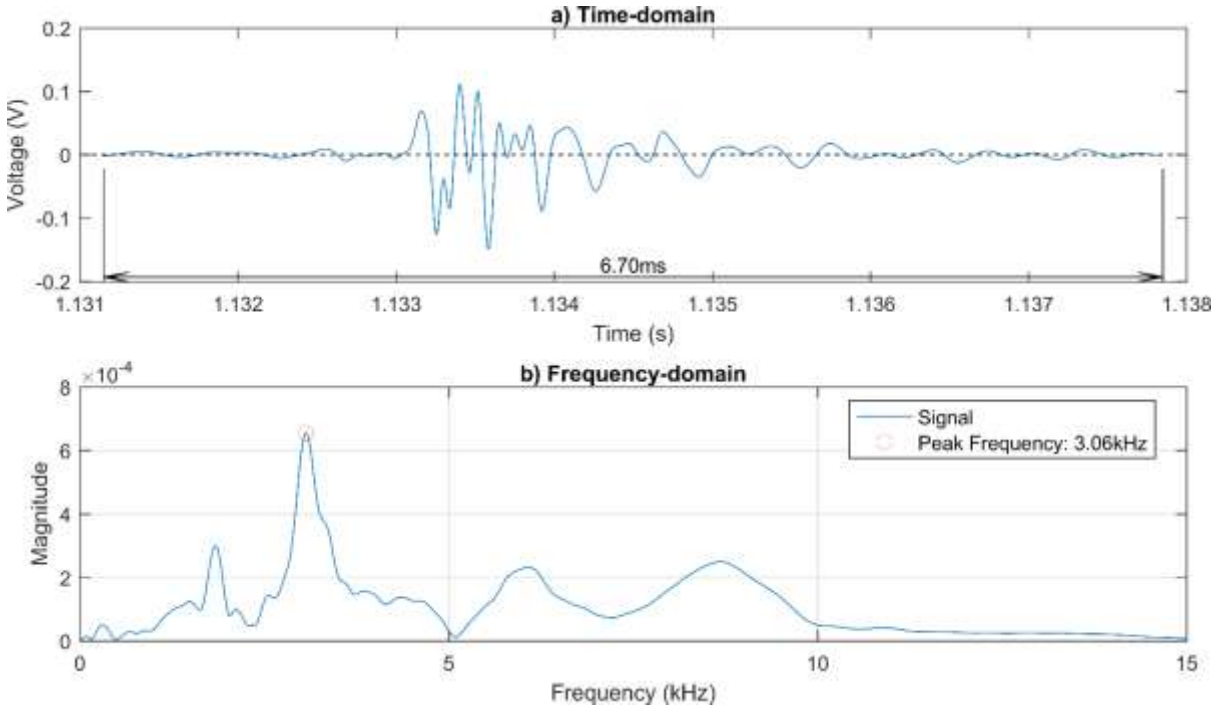

Figure 4.10: (a) Time-domain and (b) frequency-domain response plots of an in-vivo AE event from Participant 1, recording 3.

Figure 4.11 shows the time and frequency domain responses of an *in-vivo* short duration event from a different participant (Participant 112). It was observed that the behaviour of the signal was similar to that of Figure 4.10 in that the signal indicates an impulse excitation has occurred. The most notable feature of the signal is that it looks very similar to a decaying sine wave with only one frequency component. The frequency spectrum (Figure 4.11b) supported this feature as it showed the presence of a single relatively narrow bandwidth frequency magnitude peak at approximately 3.1 kHz dominating the frequency response of the signal. A much lower magnitude and broader peak was present at approximately 1.1 kHz but the effect of this was not obvious in the time domain response of Figure 4.11a.

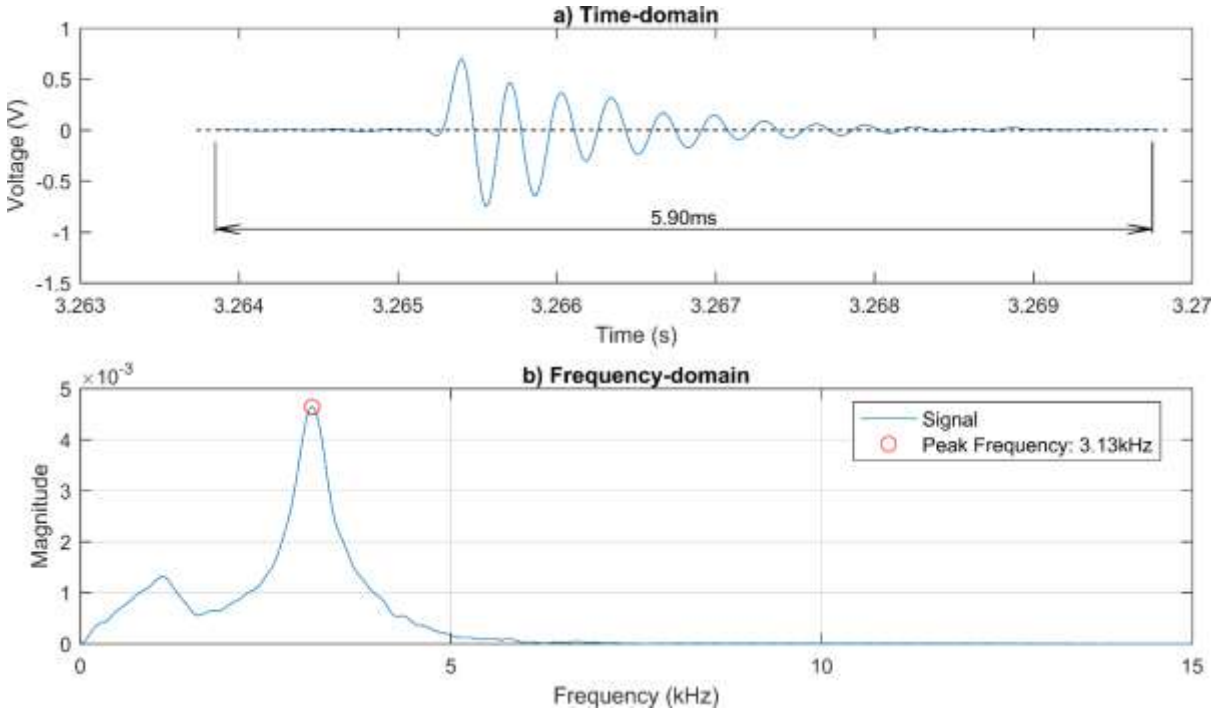

Figure 4.11: (a) Time-domain and (b) frequency-domain response plots of an in-vivo AE event from Participant 112, recording 26.

The signals of Figure 4.10 and Figure 4.11 were typical of *in-vivo* short duration AE events that occurred when long duration events did not occur in close time proximity to the short duration event. However, a number of short duration events did occur in close proximity to the long duration squeaking events. It was noted in Section 4.3.2 that some squeaking events had distinct pulses that occurred. This meant the signal returned to background noise levels for a brief time (typically only a few milliseconds) before exhibiting oscillations of significant magnitude again during squeaking events. Figure 4.12 shows the time and frequency domain response of an acoustic event that occurred within a few milliseconds after the end of the squeaking event shown in Figure 4.4. Note that the event detection algorithm determined the squeaking event of Figure 4.4 to be between 2.2-2.4 seconds of the AE recording. Furthermore, the two distinct pulses that occurred after 2.4 seconds in Figure 4.4 were determined to be two separate AE events by the event detection algorithm and it is the first of these pulses that is shown in Figure 4.12.

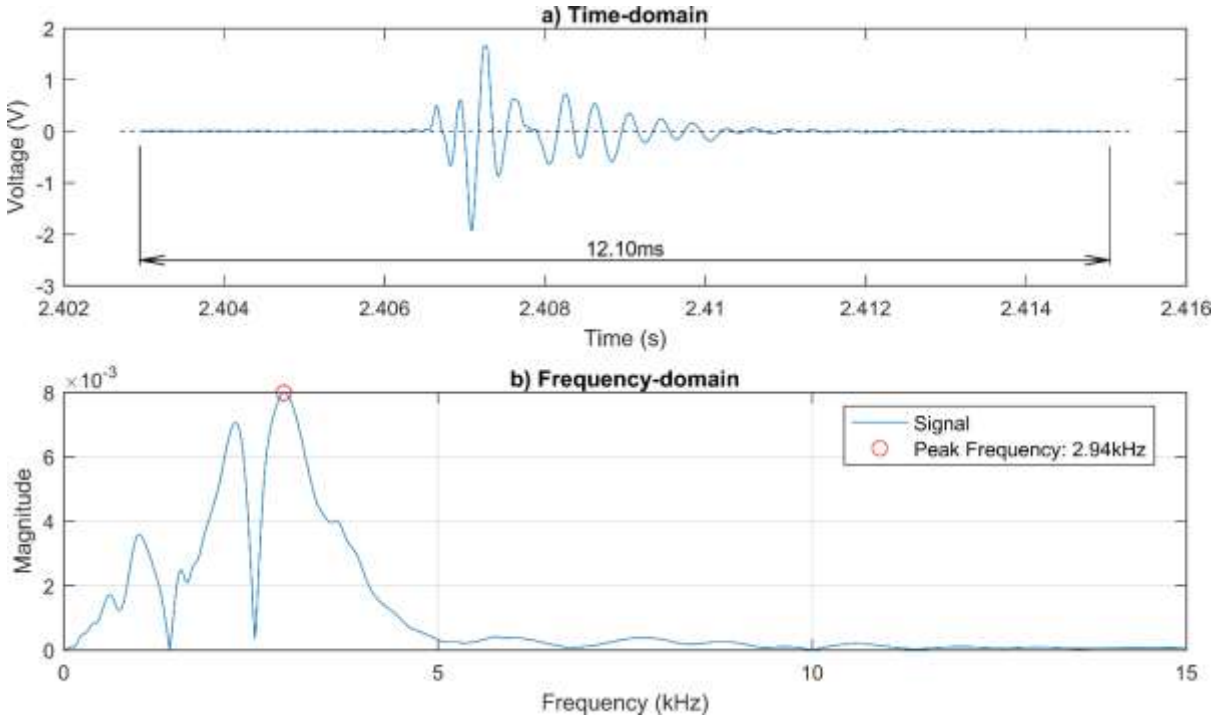

Figure 4.12: (a) Time-domain and (b) frequency-domain response plots of a second in-vivo AE event from Participant 27, recording 1, where the immediately preceding event was previously presented in Figure 4.4.

It was observed from Figure 4.12 that the acoustic event showed similar behaviour to the short duration events already presented. Nevertheless, given the close proximity of the event of Figure 4.12 to the other pulses of the audible squeak of Figure 4.4 and the overall similarity of their frequency profiles, the event of Figure 4.12 was likely to have been an audible portion of the distinct audible squeak of Figure 4.4. To the human ear, the event of Figure 4.12 sounds like a "click" when listened to in isolation. However, when the event was listened to immediately following the squeaking event of Figure 4.4 the "click" is indistinguishable from the squeaking sound and demonstrates that the event likely contributed to the squeaking sound that was observed.

Occasionally *in-vivo* short duration acoustic events occurred that had most of their frequency content above approximately 5 kHz. Figure 4.13 shows an example of such an event where it was observed from the frequency response (Figure 4.13b) the event had a main frequency of approximately 9.1 kHz and many other peaks of relatively large magnitudes at frequencies above 5 kHz. It was also observed from Figure 4.13b that the magnitude of the frequency

response for frequencies between approximately 1 kHz and 4 kHz (the range where the peak frequencies of all acoustic events presented thus far had occurred) were insignificant when compared to the magnitudes for frequencies above 5 kHz. This result indicates that while most of the short duration events were observed to behave similarly, some short duration events were recorded during testing that had main frequency content at much higher frequencies. Therefore, further investigation of the short duration events would be necessary to determine the significance of their differences and similarities.

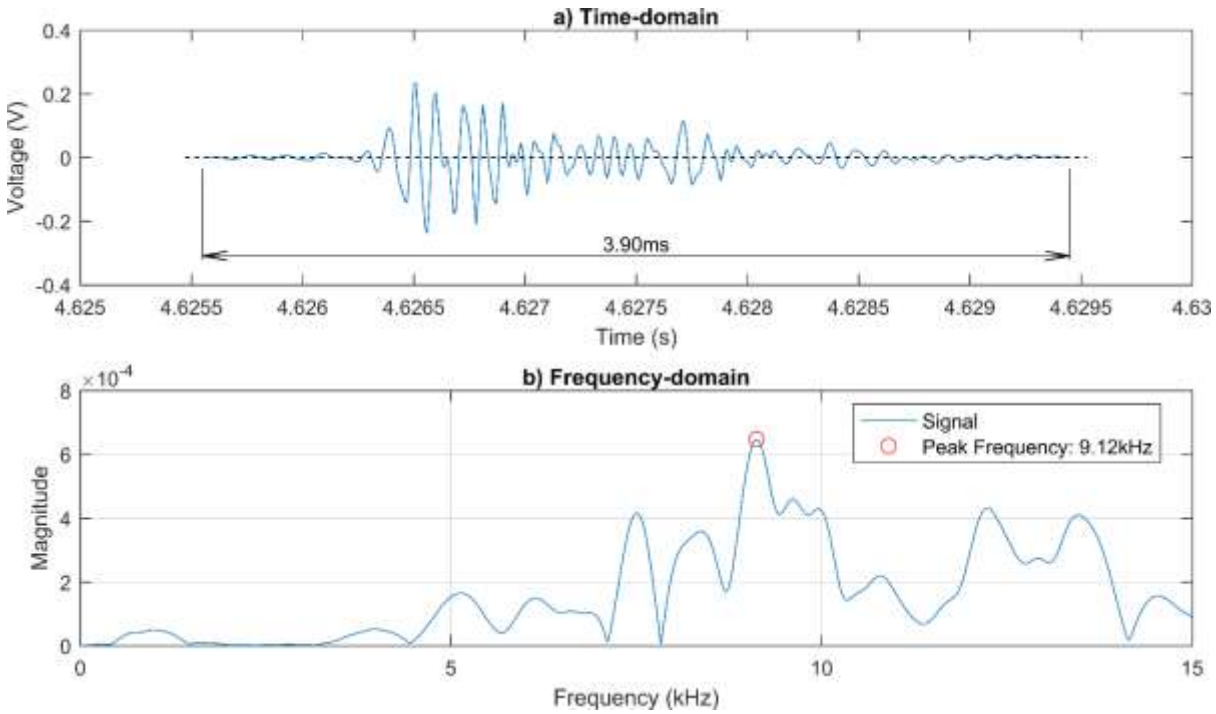

Figure 4.13: (a) Time-domain and (b) frequency-domain response plots of an in-vivo AE event from Participant 23, recording 5.

4.3.5 Control Hip Acoustic Emissions

Initial observations of the AE signals that were obtained from the control participants during *in-vivo* testing indicated a relative lack of AE activity above background noise levels (when compared to the *in-vivo* THR AE recordings). The mean rate of AE events occurring during all test recordings for the THR participants was 1.17 AE events per second, whereas the mean rate for the control participants was 0.65 AE events per second. Consequently, yielding an

approximate 45% decrease in the rate of occurrence of AE events in the control participants compared to the THR participants. Nevertheless, the total number of AEs recorded from the control participants was substantial. Figure 4.14 and Figure 4.15 show the time and frequency domain responses of an AE event that occurred during the AE recordings for two select control participants (Control 1 and Control 20 respectively). The AE events of Figure 4.14 and Figure 4.15 were typical of most AE events observed from the control participants. It was also observed that these typical events were similar in nature to the THR participant *in-vivo* short duration AE events of Figure 4.10 and Figure 4.11. This result is significant in that it indicates the high likelihood that the THR *in-vivo* AE recordings contained AE events that did not originate from the implant components. AE events that have not originated from the implant components are most likely to have been induced from sources such as other joints in the body, muscle movements, movement or sliding of tendons, fascia, ligaments, and other inherent movement of biological material. However, it is also possible that some AEs could be artefacts from the sensing hardware such as movement of the sensor array on the skin surface.

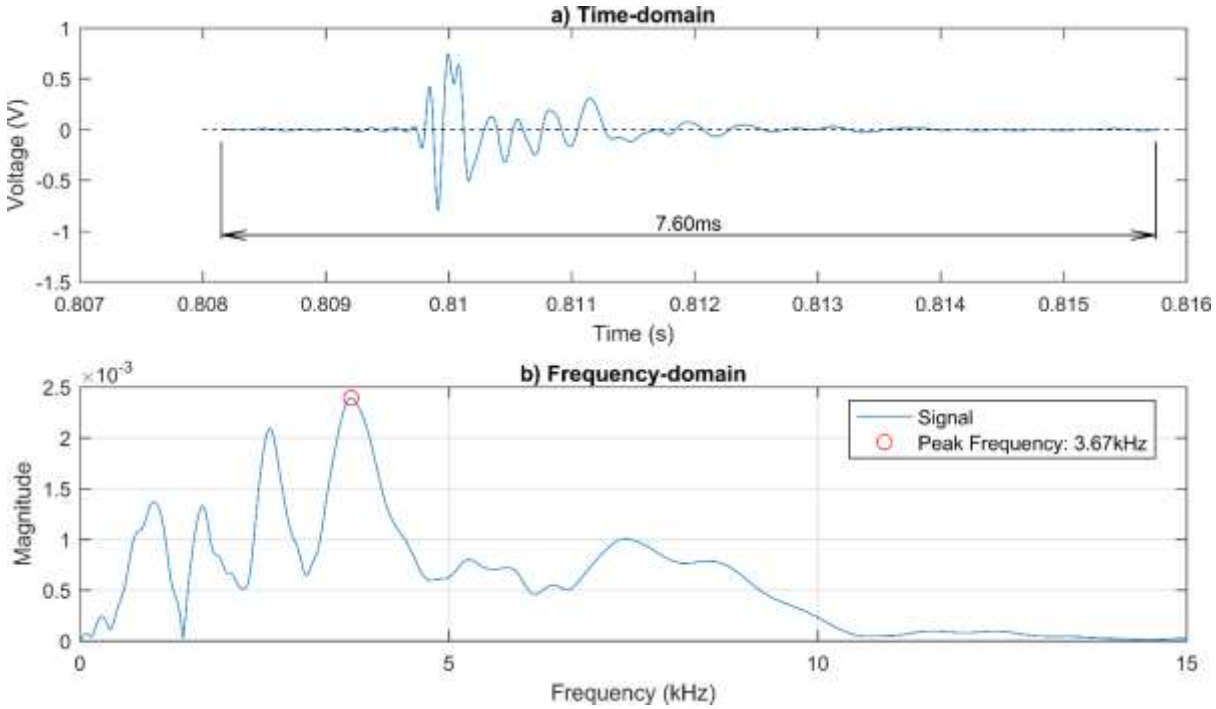

Figure 4.14: (a) Time domain and (b) frequency domain response plots of an in-vivo AE event from Control 1, recording 7.

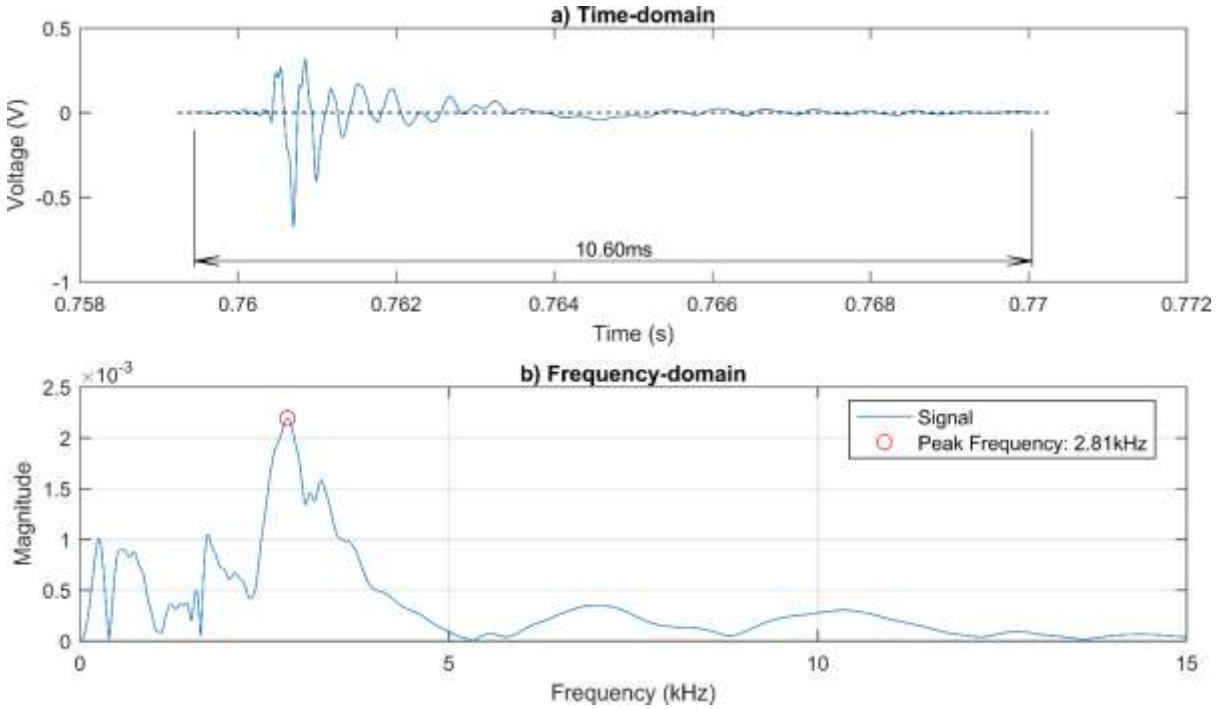

Figure 4.15: (a) Time domain and (b) frequency domain response plots of an in-vivo AE event from Control 20, recording 6.

As observed from the empirical CDFs of Figure 4.1, the control participant AE data demonstrated broadly similar trends to that of the THR participant data in which approximately 90% of the AE events had main frequency content below 3.4 kHz. Accordingly, control AE signals occasionally showed the presence of short duration AE events with main frequency content above 3.4 kHz. Figure 4.16 shows the time and frequency domain responses of a second *in-vivo* AE event from one of the control participants (Control 1). It was observed from Figure 4.16b that the main frequency content of the event was approximately 21 kHz. As with the THR participant AE data, further investigation of the short duration events from control participants would be necessary to determine the significance of the high main frequency content AE events.

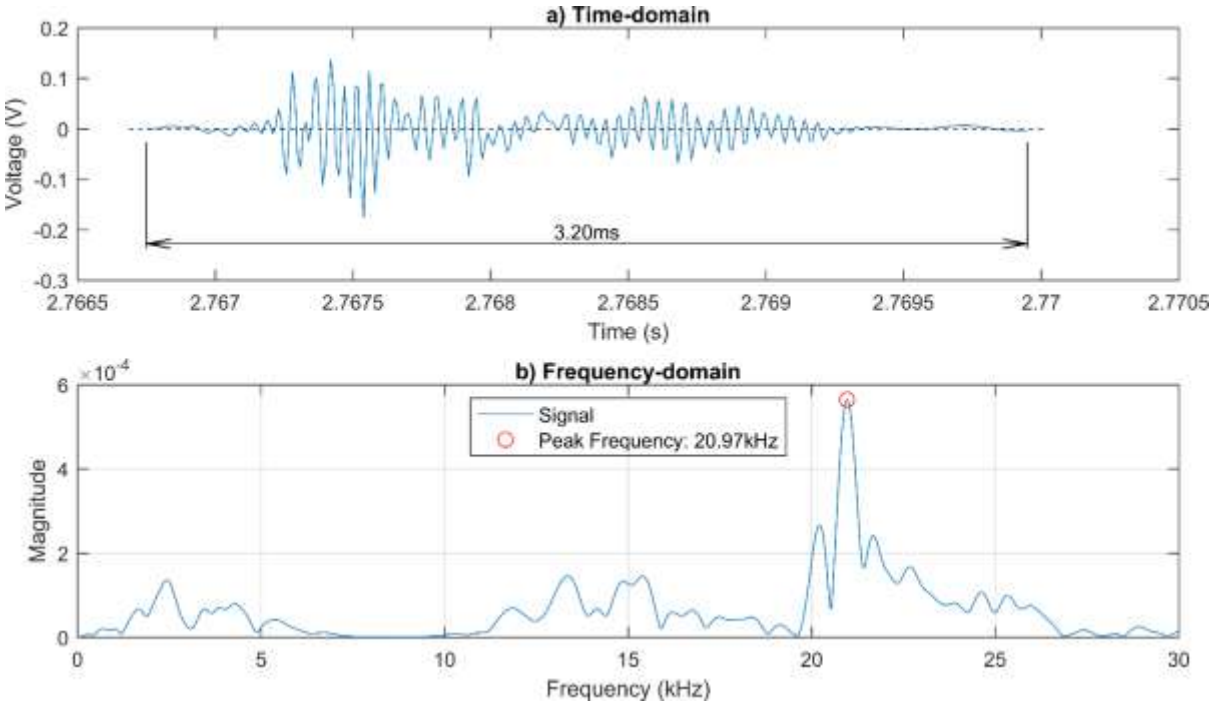

Figure 4.16: (a) Time domain and (b) frequency domain response plots of an in-vivo AE event from Control 1, recording 10.

Another observation from the control participant AE data was the presence of events with relatively long durations. Figure 4.17 shows the time and frequency domain responses of an *in-vivo* AE event from a third select control participant (Control 9). It was observed from Figure 4.17 that the duration of the AE event was approximately 40 ms, which was considerably longer than the vast majority of control participant AE events. Approximately 90% of control participant AE events had durations of 12 ms or less. Conversely, a small number of the control participant AE events had durations considerably larger than 12 ms. The typical frequency response of these long duration (>12 ms) control participant events was relatively noisy and generally contained a broad magnitude peak in the region between approximately 1 to 5 kHz as was observed in Figure 4.17. The behaviour observed from the long duration AEs in the control participant data was in contrast to the general behaviour observed in the long duration THR participant AEs of Section 4.3.2 which were generally squeaking events.

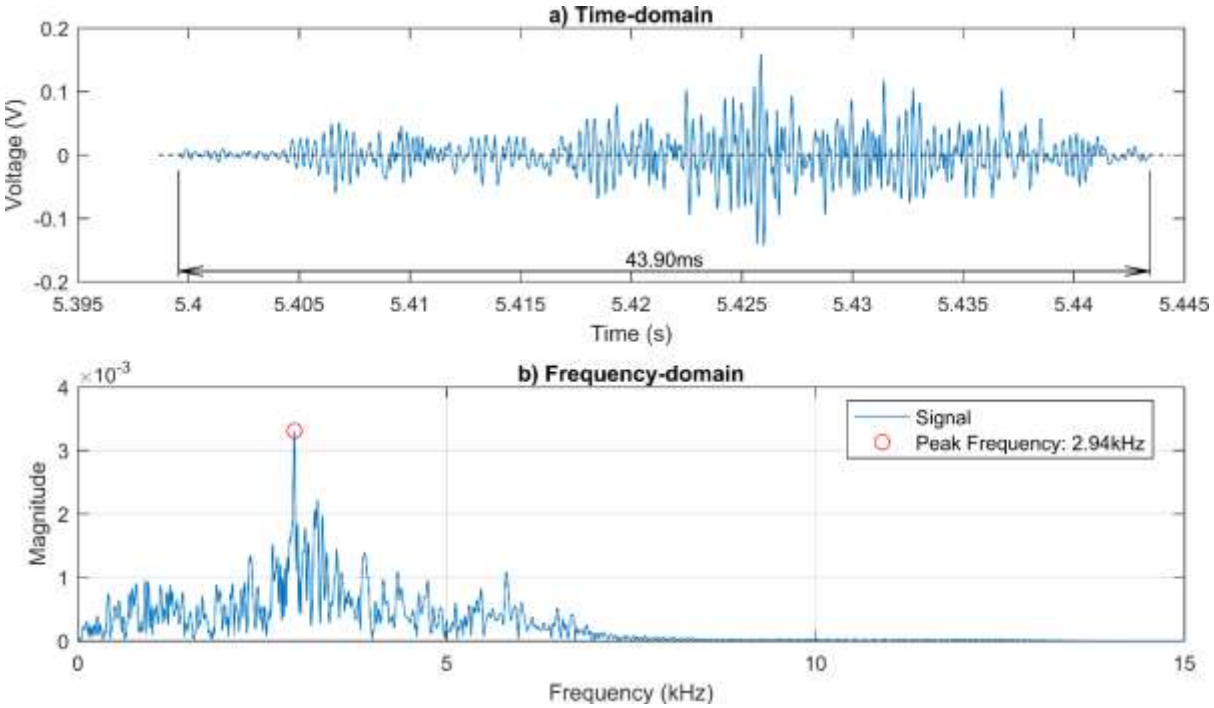

Figure 4.17: (a) Time domain and (b) frequency domain response plots of an in-vivo AE event from Control 9, recording 22.

The AEs that were detected from the control participants confirm that, as would be expected, some AEs are produced from sources unrelated to the components of a THR implant. Further analysis of this AE data from the control participants could lead to the ability to distinguish between AEs produced as a result of the presence of a THR implant and AEs produced due to interactions of biological matter unrelated to THR implants. The ability to distinguish between AEs in the aforementioned way would allow for a more focussed study on THR AEs to be carried out as AEs unrelated to the presence of the THR could be identified and excluded from analysis. Therefore, the data collected from the control participants provides useful reference AEs for the ongoing analysis of the *in-vivo* THR AEs.

4.4 In-vitro Acoustic Emissions

Initial *in-vitro* testing was focussed on ceramic-on-ceramic (CoC) implants in an attempt to replicate the audible squeaking that was produced *in-vivo* from the same implants. *In-vitro* testing of five retrieved CoC implants from the revised participants have been performed at the

time of writing. As indicated in Section 4.2, 11 participants with CoC implants had been revised. However, of these 11 participants, only five had made their retrieved implant components available for *in-vitro* testing. Four of the five available retrieved CoC implants produced audible squeaking during *in-vivo* AE monitoring confirming that the detected emissions were originating from the hip joint rather than any other biological activity. When tested *in-vitro*, under non-lubricated conditions, all five implants produced squeaking during both flexion/extension and trunnion Morse taper (TMT) rotation motions (see Section 3.5 for testing methodology).

The major observation from the *in-vitro* AE signals was that the proportion of recorded AEs that had short durations (below approximately 20 ms) was substantially less relative to the durations observed from *in-vivo* testing. Approximately 47% of *in-vitro* AE events had durations less than 20 ms, significantly contrasting the 95% from the *in-vivo* AE data. Additionally, only approximately 25% of *in-vitro* AE events had durations less than 10 ms while 80% of *in-vivo* AEs and 85% of control AEs had durations less than 10 ms. It should be kept in mind that the *in-vitro* testing was not an exact replication of the *in-vivo* testing environment. However, the *in-vitro* testing observations suggest that short duration AE events measured *in-vivo* may not be a result of interactions within the THR implant components themselves given the observed lack of short duration AEs *in-vitro*. As identified in Section 3.5, during the *in-vitro* AE monitoring an acoustic sensor was attached directly to the implant femoral stem and vibration isolation material was used to attenuate unwanted external vibrations propagating through to the implant from the robotic manipulation process. Therefore, the nature of the *in-vitro* test setup suggests that the vast majority of AEs detected *in-vitro* were the result of implant vibrations induced by the interaction of the manipulated implant components. Conversely, it is likely that a substantial proportion of the short duration AE events detected *in-vivo* are a result of interactions between other components of the overall

joint system (such as muscle, soft tissue, bone, ligaments, fascia, tendons, bone to implant interactions, or possibly even the knee joint).

As with the *in-vivo* AEs, the long duration *in-vitro* AEs were generally audible squeaking events. The time and frequency domain responses of these audible squeaking AEs exhibited similar trends to those of the *in-vivo* audible squeaking AEs of Section 4.3.2 (Figures 4.2-4.4). However, some differences were observed between the frequency domain responses of the *in-vitro* and *in-vivo* AEs. Each implant tested *in-vitro* generally exhibited different main frequency content to the other implants from the flexion/extension motion. Additionally, for some implants, the main frequency content of the AEs was found to be different if the relative rotational orientation of the implant bearing surfaces, along the rotational axes (see Figure 3.8) of the implant components, was changed between tests. Across the different implants tested *in-vitro*, the main frequency content varied between approximately 1 kHz and 8 kHz.

A small number of short duration AE events occurred during the *in-vitro* testing which were generally similar in form to the *in-vivo* short duration events of Section 4.3.4 (Figures 4.10-4.12). However, the *in-vitro* short duration events were observed to occur with more periodic consistency between their occurrences than those observed during *in-vivo* testing. The periodicity of the short duration events was primarily a result of the periodic flexion/extension motions performed by the robot arm during *in-vitro* testing, described in Section 3.5.3. The periodic consistency was also observed in the long duration *in-vitro* events and hence consistency was observed in the timing gaps between long and short duration events. Figure 4.18 shows a time domain response of AEs recorded during an *in-vitro* test where consistent timing gaps were observed between subsequent long and short duration events. The consistency in the timing gaps indicates that the occurrence of long and short duration THR AE events may not be independent of each other which may be useful information for categorising both *in-*

vitro and *in-vivo* THR AEs. While it was immediately obvious that most *in-vitro* events had this consistency in timing gaps, without further investigation, it cannot be concluded that the *in-vivo* events did not also exhibit similar consistency behaviour.

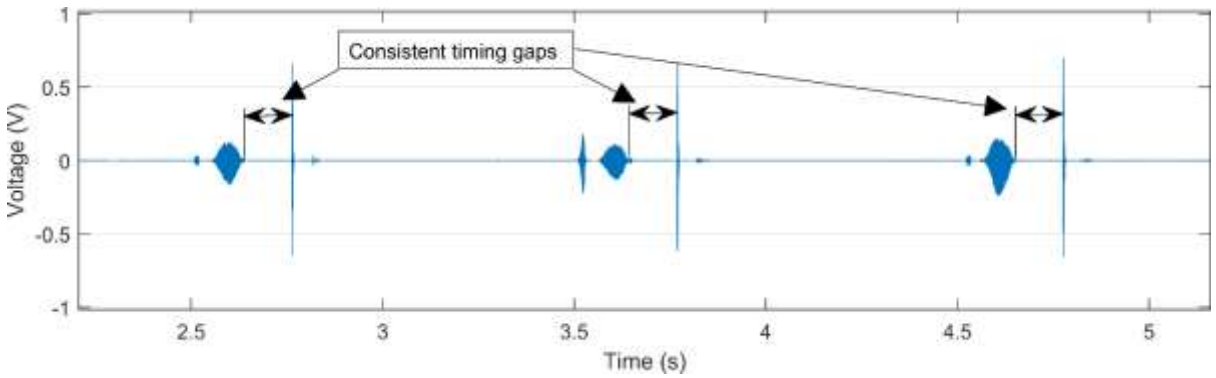

Figure 4.18: Time domain in-vitro flexion/extension test result showing consistent timing gaps between long and short duration events.

Further observations from the *in-vitro* testing were related to the AEs generated from manual trunnion Morse taper (TMT) rotation of the femoral head on the femoral stem, described in Section 3.5.4. The overall frequency responses of the *in-vitro* TMT rotations were noticeably more variable, in terms of main frequency content, and contained frequency content above background noise levels across the entire measured spectrum. Some of the variability observed in the TMT rotation AEs may be explained by inherent inconsistencies introduced by the manual manipulation of the femoral implant components. The nature of the manual manipulation meant that the force, speed, and rotation that was applied to manipulate the joint was not able to be controlled as accurately as with the flexion-extension manipulations using robotic control. Therefore, it would be desirable to conduct further TMT rotation testing to investigate the effect of different testing conditions on the observed AEs, preferably using robotic manipulation.

Nevertheless, it was observed that most AEs generated from the TMT rotation had main frequency content above approximately 12 kHz. On occasion, audible squeaking was observed during TMT rotation testing. The observed squeaks generally had a noticeably higher pitch

than the squeaks of the flexion-extension motion which was consistent with their frequency content being above 12 kHz. Additionally, occurrences of large voltage values from the ultrasonic receivers were often observed during TMT rotation testing, indicating substantial acoustic activity had occurred, but these occurrences had not been perceived by the naked ear. Given that the frequency range of human hearing only extends up to approximately 20 kHz, the unperceived AEs necessarily had their main frequency content above 20 kHz which was confirmed by examining their frequency responses. Figure 4.19 shows the time and frequency domain responses of a typical *in-vitro* AE event from the TMT rotation motion where it was observed that the main frequency content occurred at approximately 22 kHz. Across all of the *in-vitro* tests, the main frequency content of the TMT rotations generally ranged from approximately 12 kHz up to 45 kHz.

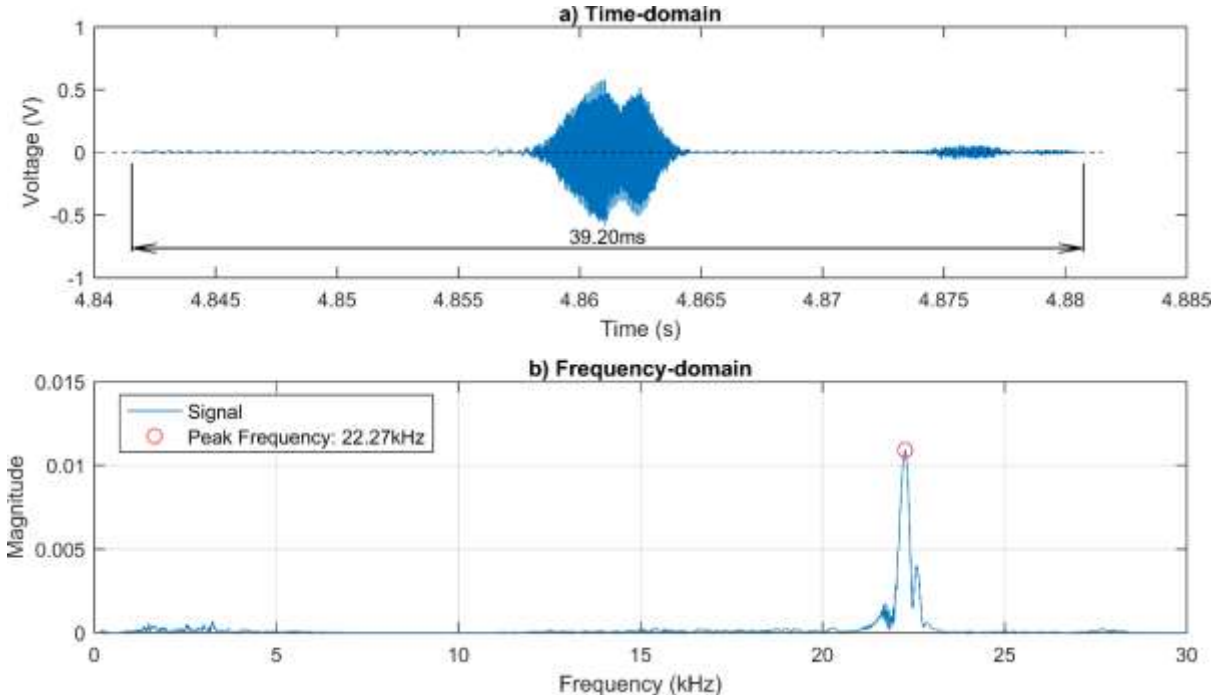

Figure 4.19: (a) Time domain and (b) frequency domain response plots of an in-vitro trunnion Morse taper rotation AE event from Participant 1 (Bench 76, recording 13).

Another important observation from the TMT rotation testing was that grinding noises were often heard as a result of friction induced by the implant interactions. Frequency responses of multiple TMT rotation tests where audible grinding occurred, but no audible squeaking, were

examined and were found to contain main frequency content in the 1-5 kHz range. Figure 4.20 shows the time and frequency domain responses of an *in-vitro* TMT rotation AE where audible grinding was observed. It was apparent from Figure 4.20 that the main frequency content of the grinding noise was at approximately 2.4 kHz. The observation that TMT rotation grinding noises have main frequency content in the range below 5 kHz is important because the audible AEs of the flexion-extension motions also had main frequency content within this range. It is also important to note that when only audible grinding noises were observed there was no noteworthy frequency content above approximately 12 kHz in the AEs.

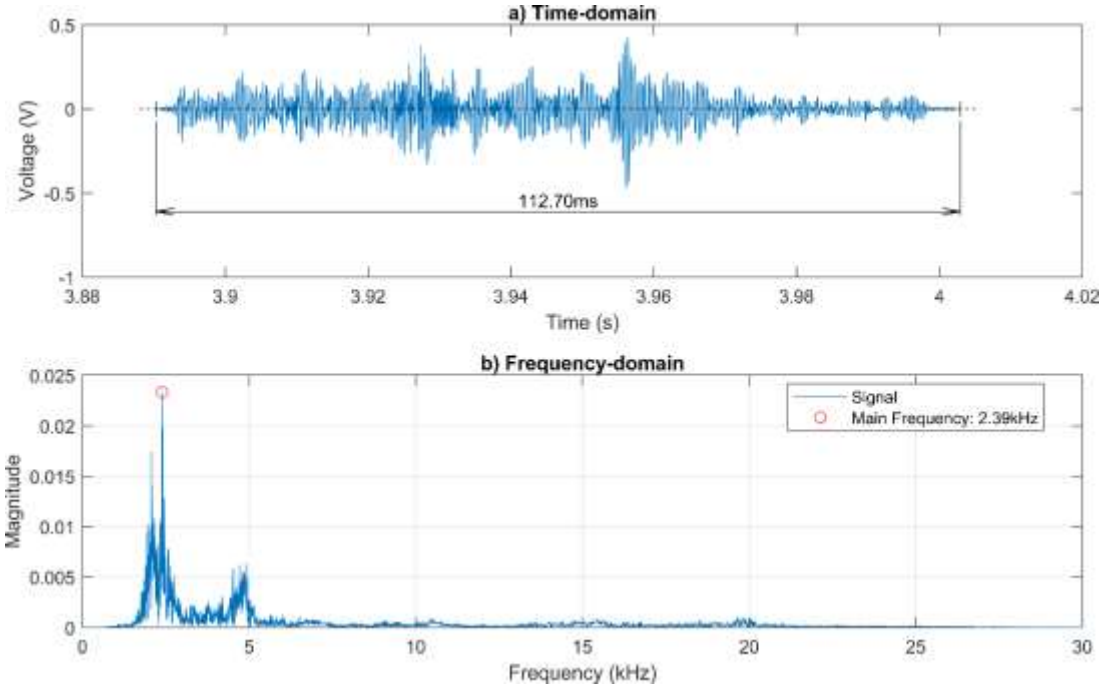

Figure 4.20: (a) Time domain and (b) frequency domain response plots of an in-vitro trunnion Morse taper rotation AE event from Participant 53 (Bench 78, recording 9).

4.5 In-vivo & In-vitro Comparisons

The frequency responses of the *in-vivo* and *in-vitro* flexion-extension motion were compared for particular participants and it was found that they exhibited similar characteristics. The frequency responses were computed using a Fast Fourier Transform (FFT). In order for the frequency domain transforms to be comparable in both frequency and magnitude, the time

domain signals were split into time windows with a duration of two seconds. A Hanning windowing function was applied to the signal data to remove any effects from non-periodicity or discontinuous signals. After the FFTs had been applied, the frequency magnitude spectrum of the windowed signals, from 0 Hz to the Nyquist frequency of 50 kHz, was plotted.

Figure 4.21 shows all the frequency responses of the *in-vivo* and *in-vitro* flexion-extension and *in-vitro* TMT rotation motions for one of the five CoC THR participants whose implant underwent both *in-vivo* and *in-vitro* testing. That is, the individual frequency response for each two second window of time has been overlaid onto the same plot and coloured corresponding to its signal motion group. Consequently, the resulting plot of Figure 4.21 emphasises the common frequencies that dominated the AE signals for each of the three signal motion groups. Figures 4.22-4.25 show equivalent plots to Figure 4.21 for the other four of the five CoC THR participants whose implant underwent both *in-vivo* and *in-vitro* testing. It was observed from Figures 4.21-4.23 that the positions of the main frequency content of the *in-vivo* and *in-vitro* flexion-extension motion datasets for three of the participants were broadly similar. However, the *in-vivo* results generally showed a much broader peak as might be expected due to the higher complexity of testing and sound transmission paths which subtly affect the position of the main frequency content. Upon closer inspection, the presence of harmonics was observed in the *in-vivo* results of all five participants. Figures 4.24-4.25 demonstrate that while the main frequency content of the *in-vivo* and *in-vitro* flexion/extension motions did not align for two of the five participants, the main frequency content of these two participants' *in-vitro* flexion-extension responses approximately aligned with a harmonic peak of the corresponding *in-vivo* responses.

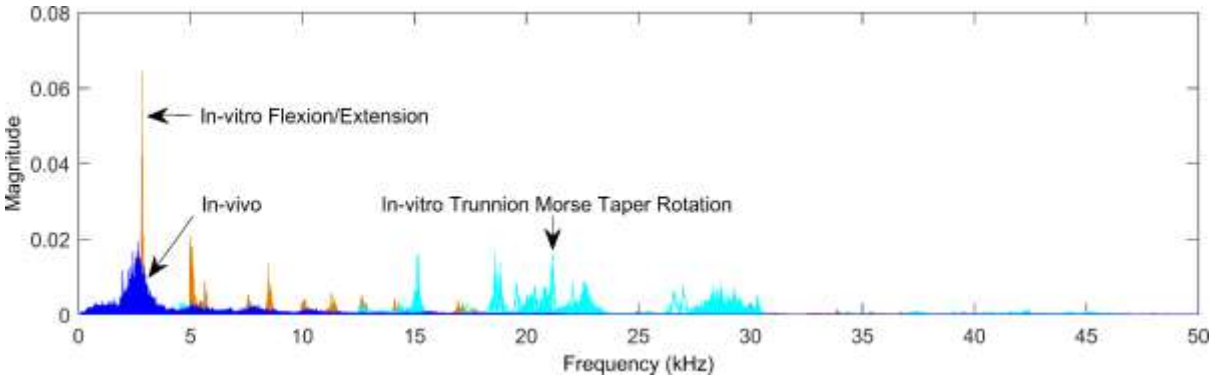

Figure 4.21: Overlaid frequency responses for all in-vivo and in-vitro flexion motions and in-vitro trunnion Morse taper rotations for Participant27.

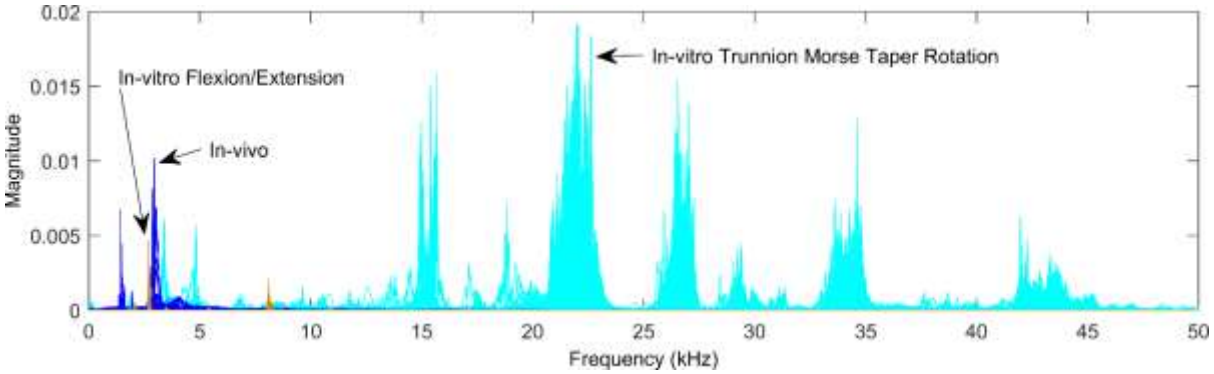

Figure 4.22: Overlaid frequency responses for all in-vivo and in-vitro flexion motions and in-vitro trunnion Morse taper rotations for Participant25.

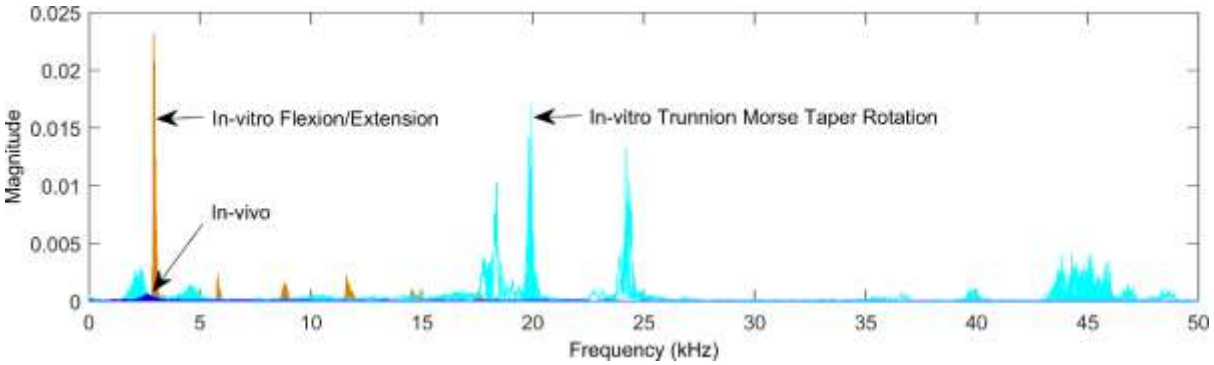

Figure 4.23: Overlaid frequency responses for all in-vivo and in-vitro flexion motions and in-vitro trunnion Morse taper rotations for a fifth select participant Participant53.

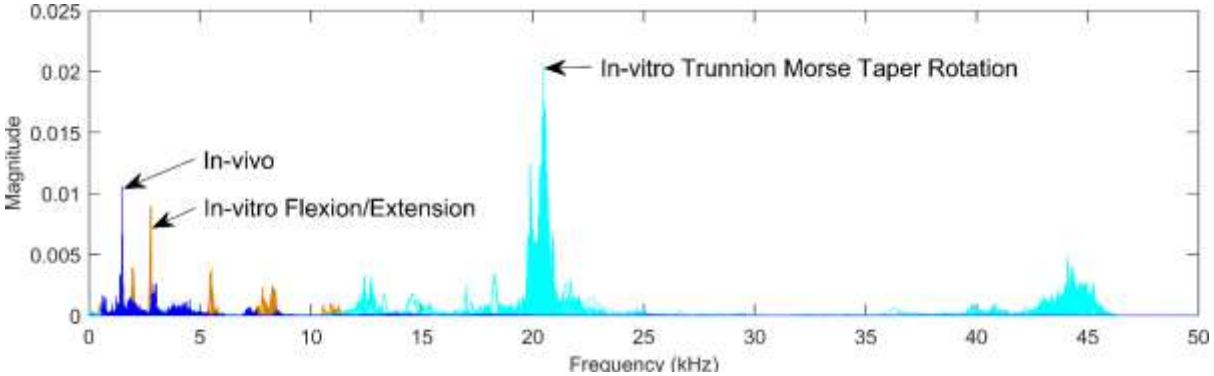

Figure 4.24: Overlaid frequency responses for all in-vivo and in-vitro flexion motions and in-vitro trunnion Morse taper rotations for Participant1.

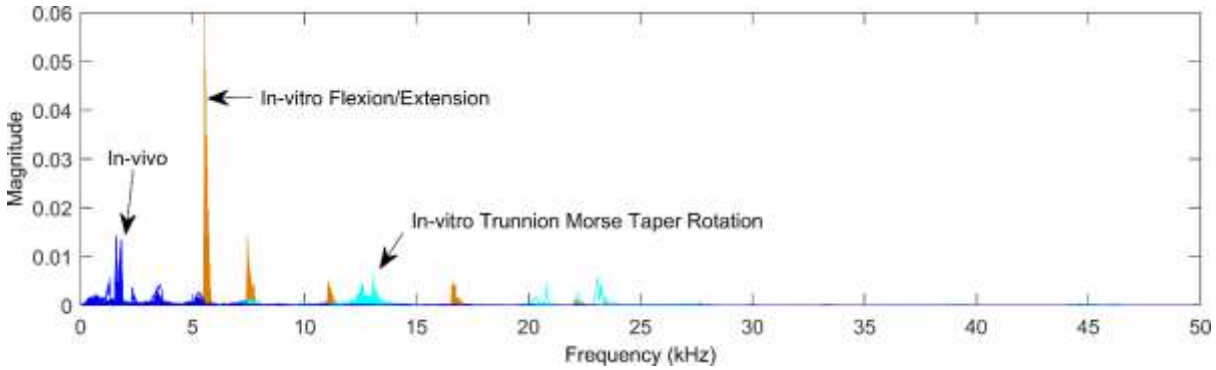

Figure 4.25: Overlaid frequency responses for all in-vivo and in-vitro flexion motions and in-vitro trunnion Morse taper rotations for Participant6.

In contrast to the flexion-extension motion AEs, the *in-vitro* TMT rotation AEs yielded main frequency content above 12 kHz as identified in Section 4.4. It was clearly observed from Figures 4.21-4.25 that the overall frequency responses of the TMT rotation AEs had more high frequency content than the flexion-extension motion AEs. The occurrence of more high frequency content in the TMT rotation AEs indicates that the vibrations induced by the TMT rotations were noticeably different to those induced by both the *in-vivo* and *in-vitro* flexion-extension motions. The frequency responses of the *in-vivo* AEs show more similarity to the *in-vitro* flexion-extension motion AEs than they do to the *in-vitro* TMT rotation AEs. Hence, the results of Figures 4.21-4.25 suggest that the AEs observed *in-vivo* are more likely to be induced by the interaction of the main bearing surfaces of the THR implant than the rotation of the femoral head relative to the femoral stem (i.e. TMT rotation). However, it was also observed that some of the TMT rotation recordings did contain AEs with noticeable frequency content below approximately 12 kHz which overlapped with the main frequency content of the flexion-extension motion AEs (see Figures 4.22 and 4.23 in particular). Further investigation determined that the frequency content that overlapped with that of the flexion-extension motion AEs was a result of a friction grinding noise at the TMT interface during the component manipulations, such as that shown in Figure 4.20, and was not the result of any squeaking AEs.

4.6 High Frequency In-vivo Acoustic Emissions

The *in-vivo* and *in-vitro* comparisons showed that the *in-vivo* AEs did not have any substantial frequency content above approximately 10 kHz. Furthermore, research into attenuation characteristics of the soft tissue surrounding THRs had suggested that there is significant attenuation of vibrations above 10 kHz (Khan-Edmundson et al., 2012). Therefore, it could be argued that the soft tissues may have masked the occurrence of high frequency AEs during the *in-vivo* testing. Consequently, comparisons were performed between ambient noise and clearly audible *in-vivo* squeaking events to reinforce the ability of the current AE monitoring equipment to detect the presence of high frequency *in-vivo* AEs. An example of the time domain response of an *in-vivo* AE recording in one participant with three audible squeaks (at approximately 0.75, 1.25, and 2.25 seconds) is shown in Figure 4.26. A spectrogram of this recording was taken to show the transient frequency content throughout the recording (Figure 4.27). It was observed from Figure 4.27 that during the three audible squeaks there was a noticeable increase in the frequency content across the entire 0-50 kHz spectrum relative to the ambient noise level during the other parts of the recording. This result is important since it confirms with good confidence that if high frequency AEs occurred during *in-vivo* testing they would have been detected by the AE monitoring equipment.

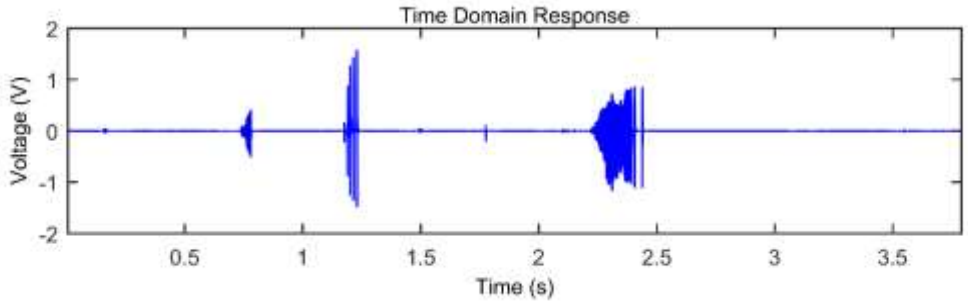

Figure 4.26: Time domain response of three in-vivo audible squeaks from Participant 1.

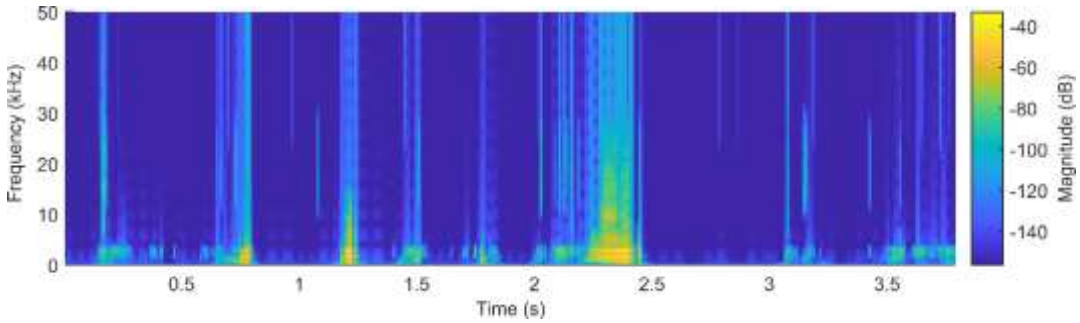

Figure 4.27: Spectrogram of time domain response shown in Figure 4.26.

4.7 Clinical Relevance of In-vivo & In-vitro Comparisons

In-vitro rotation of the trunnion Morse taper connection (TMT) produced characteristic squeaking with frequency responses which did not overlap with the frequencies observed during *in-vivo* squeaking. The *in-vitro* flexion-extension (FE) results, representing AEs from the main bearing surface interaction, showed a notably better fit to the *in-vivo* data than the *in-vitro* TMT rotation results. In spite of directly observed fretting of the trunnion Morse taper, this would strongly suggest that the AEs generated from a participant's hip prosthesis are likely to be induced from the main bearing surface interaction rather than any rotation occurring at the TMT.

As mentioned in Section 4.6, previous research had suggested that significant attenuation of vibrations above 10 kHz occurs due to the soft tissue surrounding a THR (Khan-Edmundson et al., 2012). It can be reasoned that AEs seen *in-vivo* may in fact still have contributions from TMT rotation as only low frequencies reach the sensors on the skin's surface. However, it was

demonstrated that the AE sensors recorded an increase in the frequency content across the entire 50 kHz spectrum relative to ambient noise during *in-vivo* testing (Figure 4.27). This implies that high frequency (>10 kHz) AEs would in fact reach the skin and be detected by the sensors, albeit at reduced magnitude. From the results, it can be concluded that squeaking *in-vivo* is not likely caused by TMT rotation as the presence of characteristic high frequency TMT rotation squeaking AEs was not detected in the *in-vivo* tests of the 111 participants in this study. While it is true that a very small number of *in-vivo* recordings did have main frequency content above 10 kHz, the durations of these high frequency AEs were too short (i.e. less than 10 ms in duration) to be observed as squeaking.

These findings suggest there is no significant movement occurring at the TMT *in-vivo* and, by implication, we should not expect to see excessive wear on this interface. However, wear at this interface is extremely common and was shown on the retrieved implants within this study and on those examined by others (Goldberg et al., 2002; Hallab et al., 2004; Hannouche et al., 2010; Jacobs et al., 2014; Kop & Swarts, 2009). Figure 4.28 shows an image of a retrieved femoral head from this study with obvious wear markings in the Morse taper region (see arrow) of the component. All five of the retrieved CoC components in this study had such surface markings in their Morse taper region.

Figure 4.28: Retrieved thirty-six millimetre ceramic femoral head showing wear markings on the Morse taper (see arrow).

This TMT wear is believed to be due to micro-movement and micro-separation as synovial fluid penetrates the junction to initiate fretting corrosion and metallosis (Hallab et al., 2004; Kop & Swarts, 2009). The larger the femoral head size, the greater the extent of trunnionosis that can occur, with subsequent metal ion release (Smith et al., 2014). The five key participants in this study all had moderately sized implants (32 mm and 36 mm). Micro-separation has also been demonstrated to occur between the ceramic head and liner in certain movements which may contribute to wear (Nevelos et al., 2000). The production of debris from this wear may then influence function of the main bearing surface. Furthermore, fretting of a ceramic head on a metal neck has been shown to occur less often than with a metal-on-metal connection (Hallab et al., 2004). Proper application of the head onto the neck at time of surgery is the most important step to reducing the prevalence of such wear.

These observations support the conclusion that squeaking originates from the bearing surface secondary to lubrication film breakdown and third body debris. Chevillotte et al. (2010) observed that there was always squeaking in non-lubricated conditions and when lubrication is present, noise only occurs when metal debris is present between the femoral head and acetabular liner. Furthermore, the disruption of lubrication appears related to the presence of third body wear (i.e. particles between liner and head) (Chevillotte et al., 2010). These third body particles can be produced by neck impingement (Ecker et al., 2008; Walter et al., 2004), stripe wear formation (Taylor et al., 2007), and fretting of the TMT articulation (Goldberg et al., 2002; Jacobs et al., 2014). Furthermore, upon physical inspection of the retrieved THR implant components in the present study, visible wear was apparent in particular locations on the components. It has been reported in literature that retrieved CoC femoral heads sometimes have a long, narrow area of damage known as stripe wear; after a period of time the narrow stripe becomes a broader wear region (Walter et al., 2006). Figure 4.29 shows obvious broad stripe wear on the main bearing surface of a retrieved 36 mm femoral head and its matching

acetabular liner. Note that the components shown in Figure 4.29 had a thin layer of carbon coating applied so that their surfaces could be examine under a scanning electron microscope (SEM) and as a result the stripe wear marking became much more visible. It has been reported that the presence of stripe wear alone is not related to squeaking (Walter et al., 2004) but can accelerate its production (Chevillotte et al., 2010). Thus, the stripe is an indicator of wear but is not necessarily associated with squeaking much like the fretting observed at the TMT.

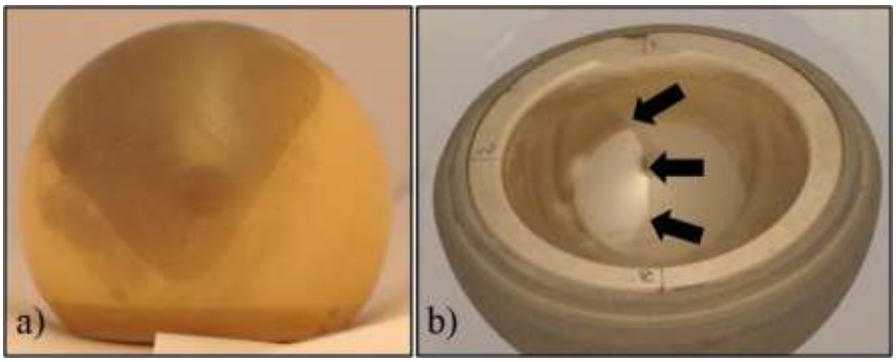

Figure 4.29: a) Retrieved thirty-six millimetre ceramic femoral head showing an obvious broad stripe wear region after carbon coating and b) matching retrieved ceramic acetabular liner also showing obvious visible wear region (with boundary indicated by arrows) after carbon coating.

Femoral neck length and rim design are important risk factors when considering the production of squeaking (Keurentjes et al., 2008; Parvizi et al., 2011). A shorter neck length reduces the range of motion and increases the risk of impingement on the acetabular cup occurring during daily activities (walking, standing from chair) (Keurentjes et al., 2008). Another factor that increases impingement is an elevated rim of the acetabular cup designed to protect the ceramic from direct contact with the femoral neck (Parvizi et al., 2011; Walter et al., 2011). As noise is not typically produced at the extremes in range of motion, and an elevated rim is not the only design-dependant feature for squeaking (Parvizi et al., 2011), it is unlikely that neck impingement itself is producing noise. Instead, the contact produces metallosis with debris acting as third bodies to disrupt the lubrication film (Ecker et al., 2008). This contributes to the multifactorial nature of debris production followed by lubrication disruption and increased friction at the bearing surface to produce noise.

4.8 Summary

This chapter has presented an initial overview of AE signal characteristics and some broad trends observed within the data. This initial presentation of key signal characteristics and overall trends was presented so that the more details results and quantitative comparisons made in subsequent chapters can be considered in context.

The results presented within this chapter indicate that, approximately 95% of the clear individual events that were identified from the *in-vivo* AEs, were found to have durations below 20 ms. The vast majority of these individual events were not identified as audible during the testing but a small number were associated with audible clicking sounds. Audible squeaking was observed from some of the *in-vivo* THR AE testing and these audible squeaking events were all found to have durations longer than 20 ms. In general, the *in-vivo* audible squeaking AEs showed much variability in terms of overall duration and the form of their time domain response. Nevertheless, the frequency responses typically had main frequency content in the range of 1-4 kHz and showed either distinct narrow-band peaks (Figure 4.3), sometimes indicating strong harmonics (Figure 4.2), or more broad-band noisy peaks (Figure 4.4). The *in-vivo* non-squeaking THR AEs were generally found to have main frequency content in the same 1-4 kHz range. However, there were a small number of non-squeaking AEs observed to have main frequency content above 4 kHz.

An important difference between the control AE observations and those of the *in-vivo* THR AEs was that the occurrence of clear AE events was noticeably less frequent in the former. That is, in terms of AEs, control hips were much less noisy than the hips of THR participants. However, *in-vivo* control AEs were commonly found to have similar behaviour to the non-squeaking THR AEs. Therefore, it was concluded that there was a high likelihood that the THR

in-vivo AE recordings contained many AE events that did not originate from the implant components.

In-vitro flexion-extension AEs were observed to have frequency responses with main frequency content similar to that of the *in-vivo* THR AEs. Conversely, the *in-vitro* TMT rotation AEs were observed to have frequency content which was at distinctly higher frequencies, typically in the range of 18-22 kHz. The *in-vitro* flexion-extension results, representing AEs from the main bearing surface interaction, showed a notably better fit to the *in-vivo* data than the *in-vitro* TMT rotation results. The results strongly suggest that the AEs generated from a participant's hip prosthesis were likely to be induced from the main bearing surface interaction rather than any rotation occurring at the TMT. Based on quantitative AE records and frequency analysis in this study, it does not appear that audible squeaking in THR patients occurs due to movement at the TMT connection, albeit from a limited *in-vivo* data set. Wear occurring at the TMT junction likely contributes to the presence of debris at the bearing surface which may be responsible for disrupting lubrication, increasing friction and noise production. Therefore, it is still possible that wear at the TMT interface could be indirectly causing, or contributing to, squeaking at the primary bearing interface.

A method for determining the fundamental frequency content of squeaking AEs was also presented in this chapter. The method, a cepstrum analysis, has the ability to isolate fundamental frequencies in signals that contain strong harmonics. The cepstrum analysis was applied to AE data of total hip replacement implants. The cepstrum analysis showed promise in identifying the primary vibration mode when strong harmonics were present, but resulted in large variation and poor identification in regions where these strong harmonics were not present. This method has promise, but it is important to carefully consider the frequency content of any signal in order to properly interpret the results of the cepstrum analysis.

Limitations of the initial stage of this study include the small number of retrieved implants available for testing, inability to obtain original acetabular shell and femoral stem, soft tissue attenuation during *in-vivo* testing and the manual rotation of the TMT. Retrieval of implants post-surgery for *in-vitro* testing has proven difficult; however, several of the *in-vivo* participant cohort are still awaiting surgery making it possible to acquire their implant components in the future. Acquisition of an entire THR after surgery would be the most ideal to study emissions from all components that contributed to the prosthesis as it was monitored *in-vivo*. However, this will be very rare and this research cannot dictate surgical decisions on what components are removed. As discussed earlier, another limitation is the attenuation effects of soft tissue during *in-vivo* recordings compared to those made *in-vitro*. Finally, removal of error borne from the unavoidable variance in manual rotation of the TMT could be gained if a mechanical system were developed that consistently rotated the components with equal pressures and speeds across different prostheses and is the focus of ongoing work.

This chapter presented a summary of the demographics of the study participants and initial comparisons from the analysis of the recorded AE data. The primary purpose of the initial comparisons of THR AEs was twofold; to introduce the basic characteristics of the AE signals, and to compare similarities in overall frequency content to help validate the usefulness of the AE monitoring technique as applied to THRs, before additional analyses were undertaken. The preliminary results from the AE monitoring device have shown that the *in-vitro* technique is useful at providing further insight into the mechanisms of acoustic emissions and wear of total hip replacement implants. The testing and analysis of the same implant components both *in-vivo* and *in-vitro* is a unique contribution of this research and provides important initial results for assessment of this sensing technology. Additionally, the initial results have shown promise in the AE monitoring technique as a diagnostic tool for the evaluation of the condition of THR implant components. Through the analysis of the time and frequency domain content of the

THR AEs, the AE monitoring technique has demonstrated that some behaviours are common, and others unique, to the AEs of different study participants.

Chapter 5 Acoustic Event Analysis

5.1 Introduction

This chapter presents the quantitative analysis of AE events. An event detection algorithm was employed to identify distinct AE events and quantify their associated signal metrics across the entire AE dataset. Due to the large number of AE signals that were recorded, it was more feasible to develop quantitative measures for comparison as opposed to qualitative visual inspections of the voltage-time signals which was inherently a manual and slow process. This chapter begins by providing an explanation and subsequent performance evaluation of the AE event detection algorithm as applied to the AE signals. The remainder of the chapter presents a quantitative analysis of the detected AE events based around comparisons of the AE event metrics. Overall statistics of the AE event metrics are presented and comparisons are drawn between the AEs from the control participants and the noise, loosening, and excessive wear revision reason participant groups. Subsequently, the results from a series of statistical hypothesis tests investigating the event metric distributions of the participant groups are presented. Finally, results from multidimensional metric comparisons are presented to gain more insight into correlations between the different AE event metrics.

5.2 Event Detection Algorithm

As stated in Section 2.4.3, an AE event is usually defined as an AE pulse that exceeds an arbitrary voltage threshold and then decays back to the noise level. Therefore, to automatically detect AE events from the recorded AE signals a voltage threshold needed to be chosen. However, due to the varied nature of the biological AE signals it was not immediately obvious what value for the voltage threshold should be chosen. Previous work had been performed

using a range of different threshold values to determine an appropriate value for the event detection. However, it was found that using a simple voltage threshold produced undesirable event detection results. It was found that larger voltage thresholds resulted in the failure to detect many AE events. Furthermore, AE events that visually appear to be single events would often be split into multiple events. It was also found that small voltage thresholds resulted in AE events being defined that had very large durations. Visual inspections of these large duration detected events would tend to suggest shorter durations would have been more appropriate. In an attempt to achieve an event detection method that distinguished AE events more reliably, previous work had suggested that applying the event detection to a variance signal computed from the original AE signal would yield more desirable results. Applying the event detection method to the variance model minimises noise and amplifies periods of intensive signal activity.

The variance model event detection algorithm had three simple steps as follows:

1) The signal was divided into short and overlapping time intervals and the signal variance was calculated for each interval.
2) An event was then identified by finding a time interval where the variance value exceed an arbitrary threshold.
3) An identified event was then isolated by determining its start and end points to occur when the statistical range of surrounding time interval variances failed to exceed a second and smaller threshold.

The first step of the event detection algorithm involved dividing the discrete AE signal into short overlapping time intervals/windows and finding the variance of the signal across each window. MATLAB was used to take a signal window of 32 samples from the beginning of the signal and the variance was then calculated for the window using MATLAB's inbuilt variance

function. The window was then shifted by 10 samples forward in time and a new variance value was calculated in the same way from the new signal window. This procedure was applied repeatedly until the entire AE signal had been processed and hence a new signal (consisting of variance values) was created with a length one-tenth of the original. The window overlap created by shifting 10 samples was used in order to create a higher time resolution variance signal to more effectively capture large voltage variations in the original signal and produce a more continuous variance signal. Figure 5.1a and Figure 5.1b show an example of a time-domain signal and its associated variance signal respectively. It should be noted that the horizontal axis of the variance signal has been converted to time and hence the signals appear the same length but in fact the variance signal has less data points. It can be seen from Figure 5.1b that the variance signal shows two events of significant amplitude and the remaining small fluctuations of the original signal do not feature.

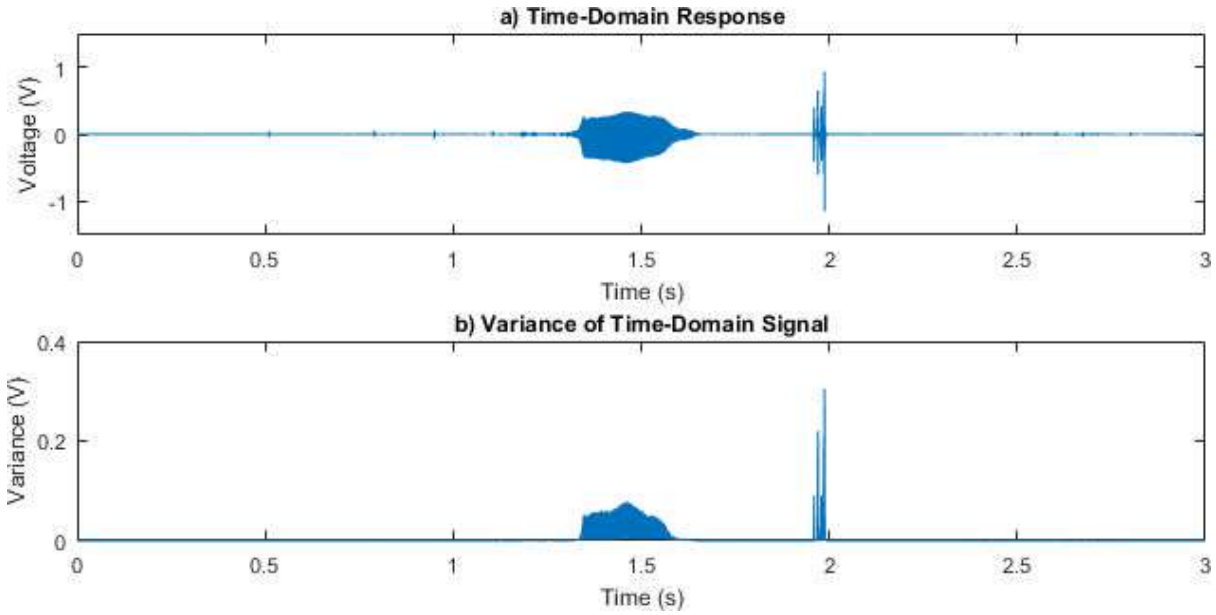

Figure 5.1: (a) An AE time-domain signal and (b) corresponding signal showing the variance.

The second step of the event detection algorithm was to use the variance signal to identify events. This was done by stepping through the variance values and checking them against a specified magnitude threshold. If a variance value was found to exceed the specified threshold

an event was identified as present and this would invoke the third step of the algorithm; the event isolation. The magnitude threshold for event identification was chosen arbitrarily to be 2.5×10^{-3} V as this value provided the most satisfactory event identification for a range of test recordings. Figure 5.2a shows a scaled version of the variance signal with the magnitude threshold overlaid. It was observed from Figure 5.2a that the chosen threshold allowed clear identification of the two perceived events of the recording whilst excluding the small duration and magnitude signal fluctuations.

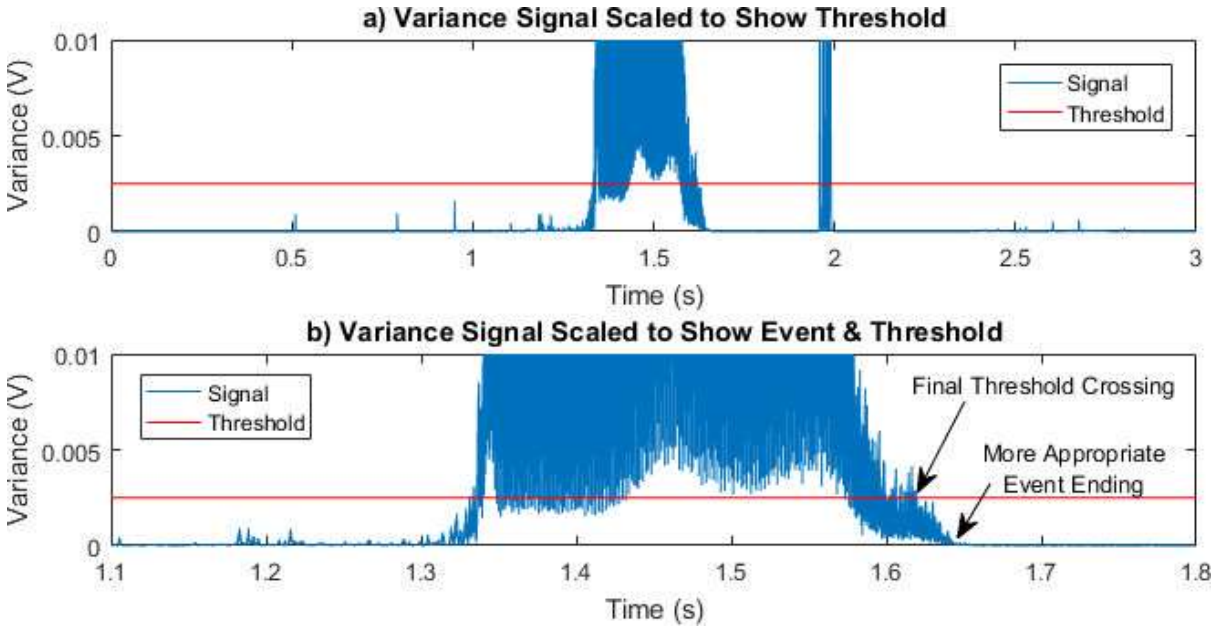

Figure 5.2: a) Variance signal scaled to show threshold and (b) variance signal scaled to show event and threshold with annotations identifying final threshold crossing and a more appropriate event ending.

However, Figure 5.2b shows a further scaled version of the variance signal and overlaid magnitude threshold and it was observed that the signal crosses the threshold a number of times during what would reasonably be defined as one event (between approximately 1.3 and 1.7 seconds) without actually returning to the background noise level. Hence, defining the beginning and end points of an event simply by threshold crossings would result in many different events being defined when in fact it may be more appropriate to define only one event. Further observation found that it may also be more appropriate for the beginning and the end of the event to be somewhat before and after the initial and final threshold crossings of an event

respectively, as is identified by the annotations in Figure 5.2b. To address these issues an event isolation algorithm was incorporated into the process.

The event isolation process occurred in conjunction with the event identification in order to prevent multiple events being identified from what would reasonably be considered a single event and also allowed more appropriate event beginning and end points to be identified. Therefore, once the identification algorithm had detected a variance threshold crossing the isolation algorithm was applied to determine the appropriate event extremities before the identification algorithm then continued from the end of the isolated event.

In order to determine the beginning of an event, the statistical range of the variance value found by the identification algorithm and the previous four variance values was taken. If the range calculated was larger than a second arbitrary threshold, known as the range threshold, a new statistical range was calculated from the five variance values shifted back by one value from the previous (i.e. add the fifth previous variance value and remove the one found by the identification algorithm). This process continued until the statistical range of the variance values became smaller than the range threshold. When the range was low enough the beginning of the event was taken to be the latest time associated with the five current variance values used for the range calculation.

Figure 5.3 demonstrates the points used for the first, second, third, and final (eleventh) range calculation of a particular AE event example. It can be observed from Figure 5.3a-c that the five points used in the range calculation step backward in time by one sample for each calculation. Figure 5.3d shows the position of the five points when the range first falls below the specified range threshold (0.5 mV for this example) and hence the event beginning point has been identified as the latest time associated with the five points used in the calculation.

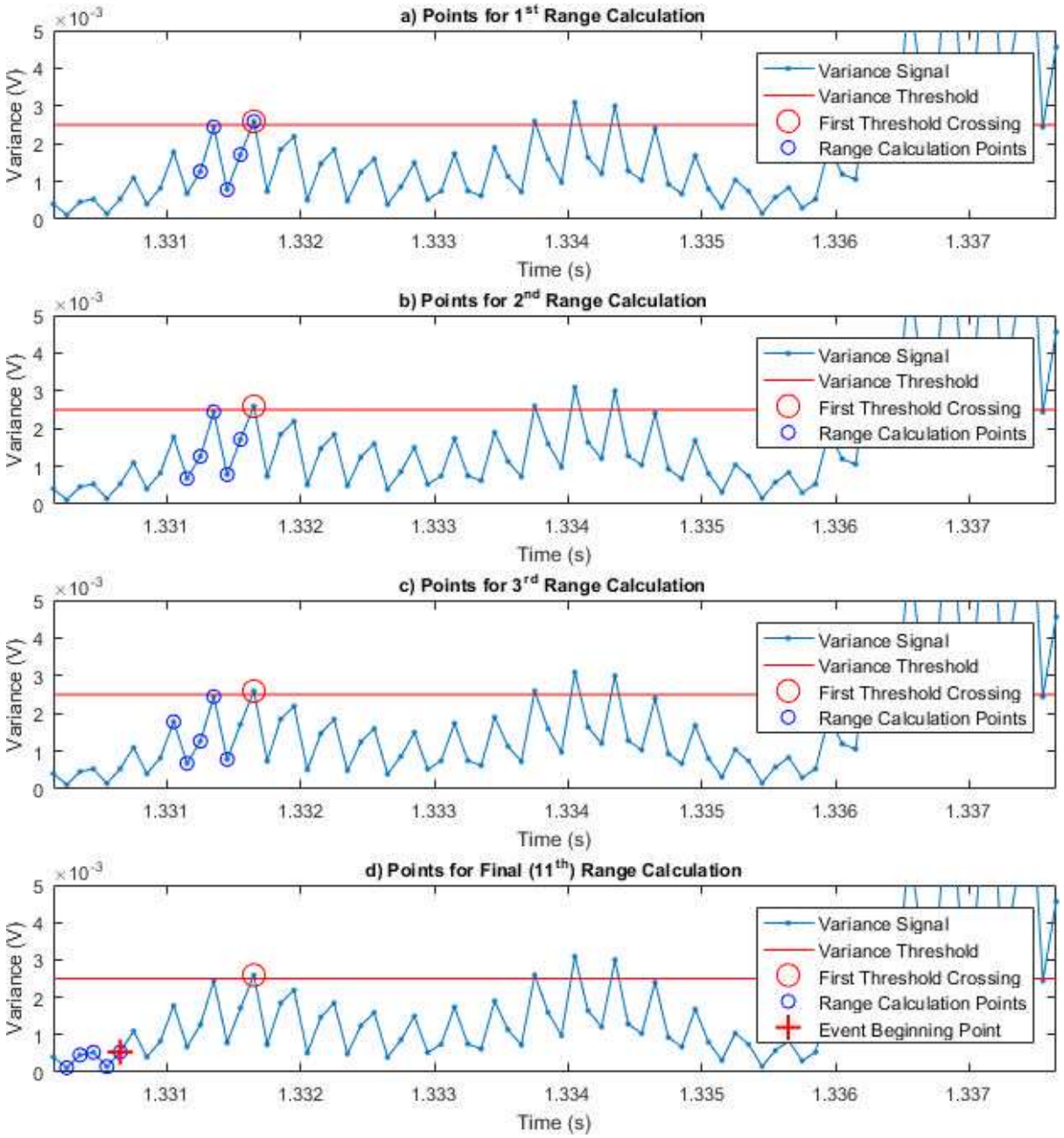

Figure 5.3: Plots demonstrating the variance value points used in the event isolation range calculation algorithm for the (a) first, (b) second, (c) third, and (d) final/11th calculations. Note that the range threshold value used to determine the event beginning point for this example signal was not the same as that which was used for the actual AE data; the threshold used here was chosen to demonstrate the algorithm more clearly.

To find the end of an event the same process was used, however, the five variance values used shifted forward in time and the end of the event was defined as the earliest time associated with the five current variance values. The value of the range threshold used for the actual implementation of the isolation algorithm was 0.005 mV and this was chosen as it produced the most satisfactory results for a range of different events.

Figure 5.4 shows the result of the event isolation (vertical lines) for the variance signal of Figure 5.2a where it was observed that the first event (Figure 5.4a) has been isolated reasonably well. Figure 5.4b shows the second event which has in fact been separated into two events by the algorithm. Hence the event detection algorithm has isolated three separate portions of interest of the input signal of Figure 5.1a and this information was then stored for later use. The effectiveness of the event detection algorithm and the chosen threshold values will be discussed in the following section.

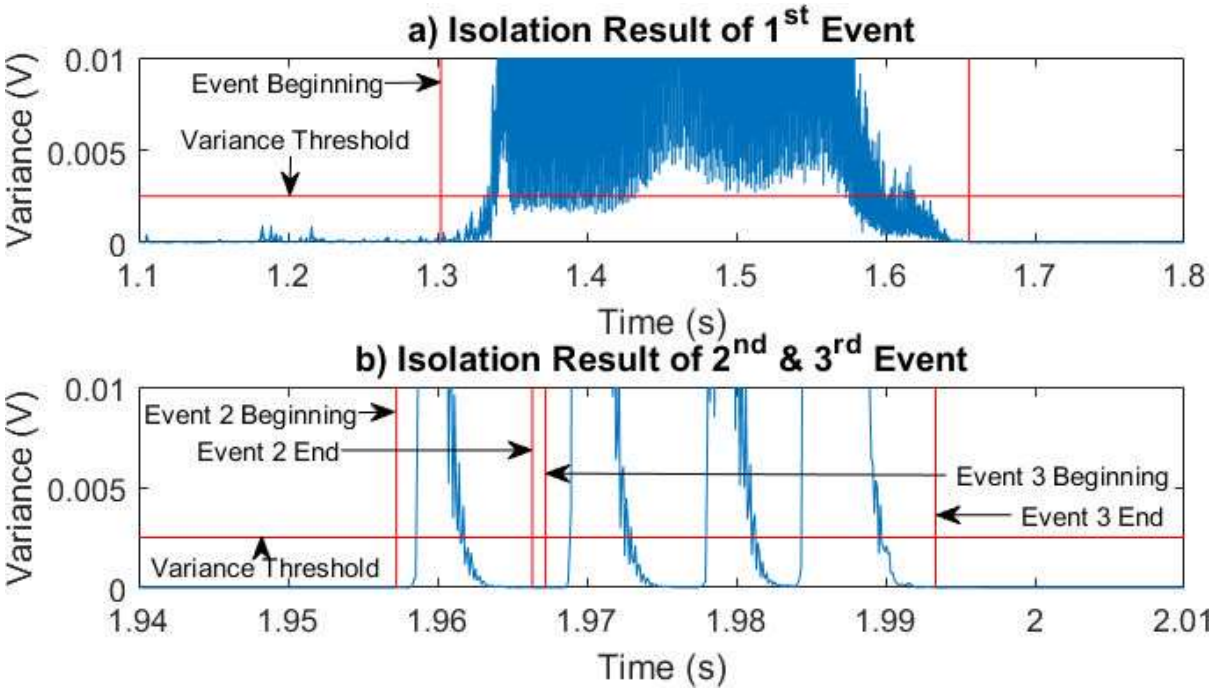

Figure 5.4: Variance signal plots showing result of event isolation algorithm for (a) first and (b) second and third events.

5.3 Event Detection Evaluation

A major limitation of the initial observations of Chapter 4 was that the frequency responses were found by computing Fast Fourier Transforms (FFTs) for two second time windows across the AE recordings. As previously identified, the AE recordings generally consisted of long durations of background noise with sporadic occurrences of AE activity above the level of background noise (i.e. acoustic events). Since the frequency responses were computed using a fixed duration window, it was certain that for some AE recordings the frequency response

obtained by the FFT was an averaged frequency response of multiple acoustic events. For example, if both a short duration event and an audible squeak occurred during a single time window the corresponding FFT would yield the average frequency response of the two seemingly unique events. The application of the event detection algorithm provided a method to address this limitation. By isolating the acoustic events the behaviours that were common and those that were unique were able to be investigated with greater rigour.

The AE recordings from the *in-vivo* data collection were typically 1-5 seconds in duration but were sometimes up to 25 seconds in length. The variation in recording durations was a result of the data collection protocol as explained in Section 3.4. Due to the relatively long durations of the AE recordings, and the intermittent nature of the occurrence of significant hip AEs, the recorded AE signals had a substantial duration of signal content that did not exceed normal background noise levels. Additionally, there were many cases where multiple unique AEs occurred during a single AE recording; a specific example of this being two separate hip squeaks occurring on two consecutive paces of a walking motion. Therefore, for further analysis it was considered necessary to separate unique AE events from single AE recordings to allow analysis of AE events separately whilst also excluding signal portions that only contained background noise.

Figure 5.5 shows examples of an AE recording for an ambient test, a control participant test, and a hip replacement participant test. The ambient test was a recording taken while the sensors were not attached to a participant and remained stationary in the test room. The control participant and hip replacement participant tests were recordings taken while the sensors were attached to a control cohort participant and hip replacement participant respectively whilst they performed a prescribed motion (squatting motion in these cases). Most notably, the ambient noise recording of Figure 5.5a remained at very low voltage amplitude for the duration of the

recording while the control hip recording (Figure 5.5b) had three major voltage spikes of very short durations and the participant recording (Figure 5.5c) had a number of voltage spikes with both short and very long durations.

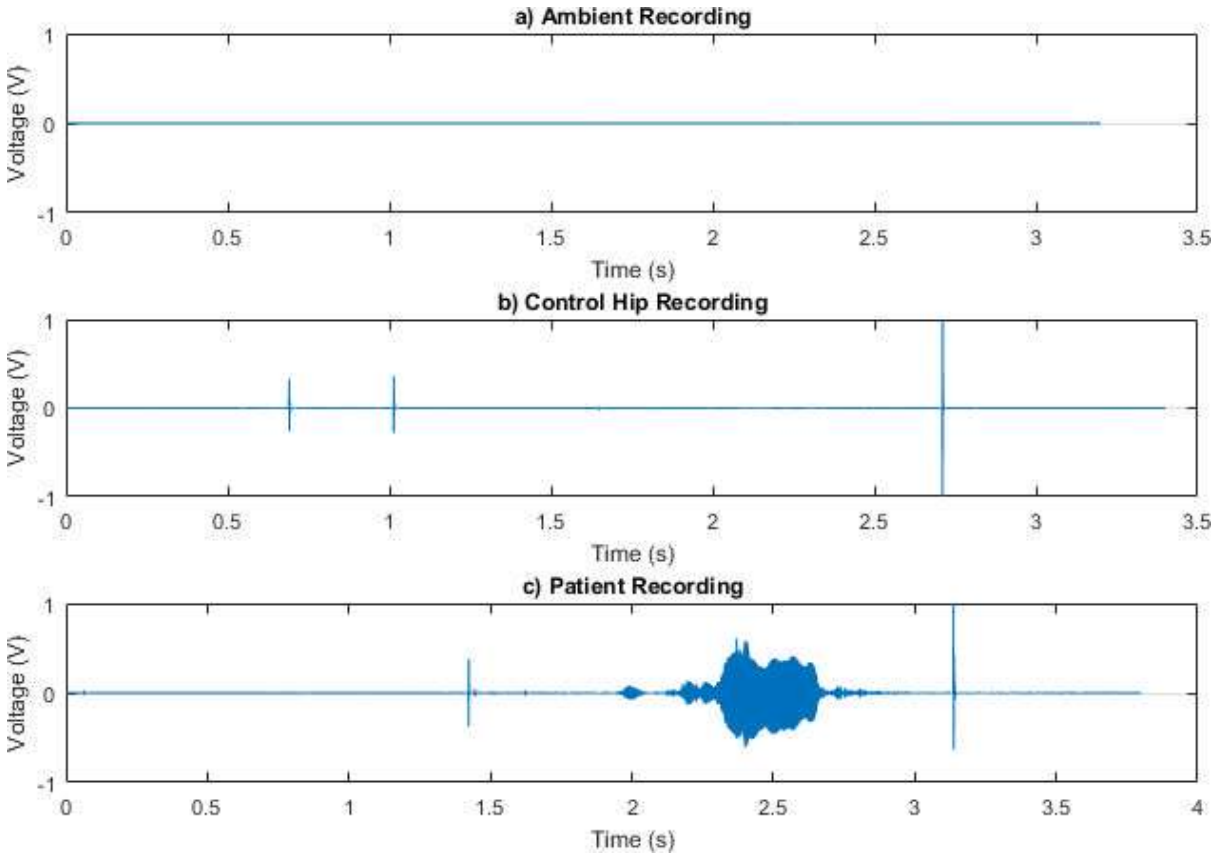

Figure 5.5: Example of AE signal recording for (a) ambient noise, (b) control hip, and (c) hip replacement participant tests.

It was evident from observing the time domain responses of AE recordings that large durations of the recordings contained very little acoustic information (i.e. regions where the voltages were consistently very low). Therefore, the regions of the recordings where voltage spikes occurred were considered to be unique AE events and separating these events from the remainder of the recordings would be useful for further analysis of the signals. Hence, the three voltage spikes of Figure 5.5b (at approximately 0.75, 1, and 2.75 seconds respectively) would be treated as separate AE events. Likewise, for the participant recording of Figure 5.5c the two short duration voltage spikes (at approximately 1.5 and 3.25 seconds respectively) would be treated as separate events and the long duration voltage spike (between approximately 2 and 3

seconds) as a third event. Additionally, the long duration voltage spike could arguably be separated into two different events as the signal does appear to return to background noise levels at approximately 2.1 seconds.

In order to perform the event separation the event detection algorithm explained in Section 5.2 was applied. Figure 5.6 shows the result of the event detection algorithm as applied to the participant recording signal of Figure 5.5c (Participant 1, recording 6). Four events were detected by the algorithm and are shown in Figure 5.6 as the regions highlighted in yellow and bordered by dotted vertical lines. As expected, the two short duration events (events one and four) were separated well. The long duration voltage spike was split into two separate events by the algorithm but the end point location of the second of these two events may not agree with that which would have otherwise been chosen by a visual inspection of the signal. A visual inspection of the signal would likely have resulted in the third event being defined between approximately 2.1 and 2.9 seconds as opposed to ending at 2.7 seconds as shown in Figure 5.6.

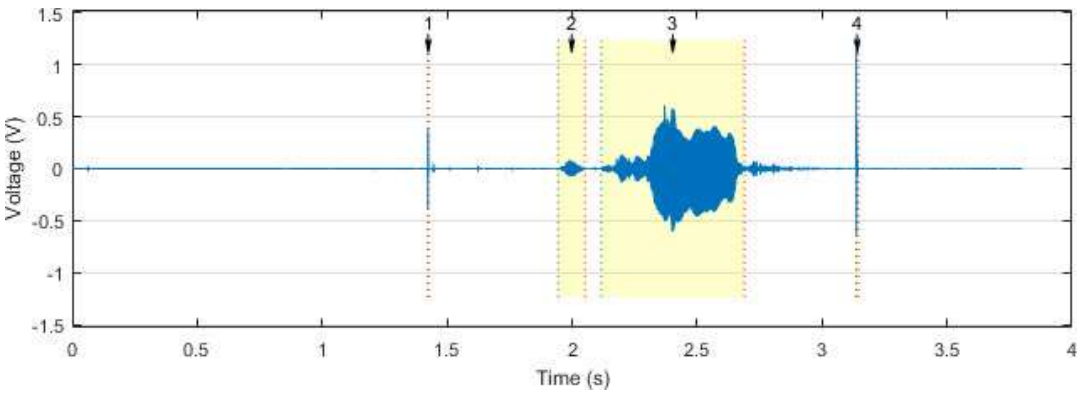

Figure 5.6: Participant AE recording from Figure 5.5c with four events, as identified by the event detection algorithm, shown bordered by dotted vertical lines and highlighted yellow.

However, it remains very subjective as to exactly where separate event start and end points should be placed. It was generally not evident whether particular AEs were from the same source; it was possible for AEs to occur from multiple sources at the same time and hence their detected signals would be superposed. Additionally, it was not clear whether two events that occurred in close proximity (with respect to time) of each other had any influence or

dependence on each other. The possibility of events influencing or depending on each other also raises the question as to whether the events should be considered together as one event anyway. For simplicity in initial analyses of this study all events were assumed to have no dependence or influence on each other. Regardless, the event start and end point decisions remain dependent on multiple factors but for the event detection process to be automated event bounds must be defined using an objective metric such as relative signal variance or amplitude threshold. Hence, the performance of any event detection algorithm will inevitably always be subjective.

Figure 5.7 shows the four events (coloured blue) of Figure 5.6, with individual axis scaling for easier viewing of the event signals, in addition to the immediately adjacent signal portions (coloured red) of each event. It was observed that event one and four (Figures 5.7a and 5.7d respectively) were isolated very well from the surrounding signal. However, the appropriateness of the event start and end points of event two and three (Figures 5.7b and 5.7c respectively) was less clear. The surrounding signals of event two and three did not appear to remain consistently at the background noise levels and had some notable signal variation that fell outside the event window.

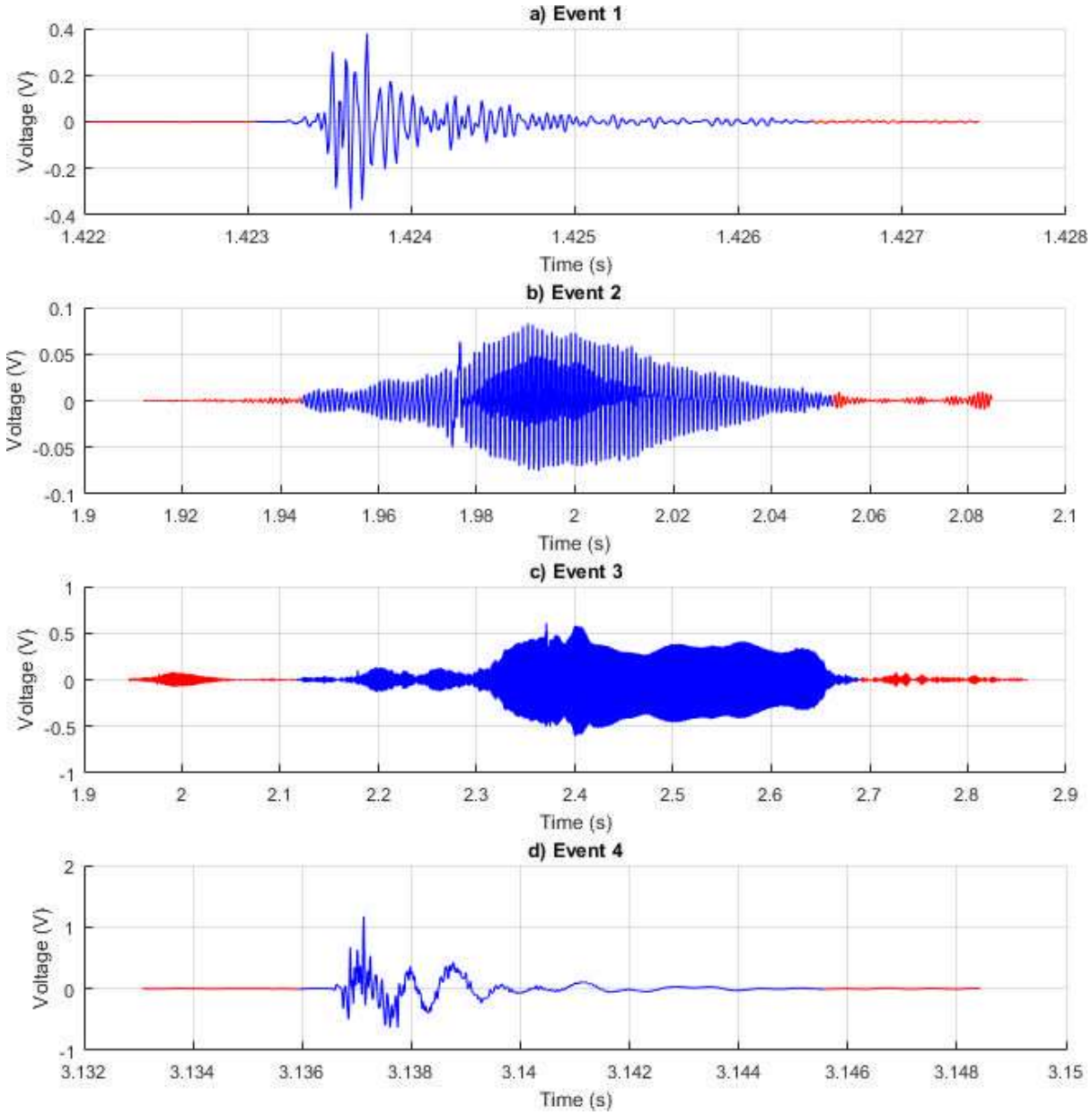

Figure 5.7: Event number (a) one, (b) two, (c) three, and (d) four from the signal of Figure 5.6. Note the event is coloured blue with the red signals being part of the original recording immediately surrounding each event.

Figure 5.8 shows a signal recording from a second select participant (Participant 27, Recording 2) and the associated events identified by the event detection algorithm. It was observed that the algorithm had identified seven separate events. However, initial observation of Figure 5.8 would seem to show that some of those events should not be separated from each other. To investigate the result from the algorithm better Figure 5.9 shows the first, second, third, and fifth events from Figure 5.8 with individual axis scaling for easier viewing. It was observed from Figure 5.9a that the first event showed an overall gradual increase in amplitude

but had intermittent decaying behaviour which was most pronounced towards the end of the event. This decaying behaviour is not evident in Figure 5.8 due to the time scale on the horizontal axis. Figure 5.9b shows the second event and it was observed that this event occurred immediately before the third, to the extent that it also appears in the plot of the third event (Figure 5.9c). From Figure 5.9c it was observed that the signal behaviour was similar to the first event in that a decay between impulses was evident but the time between the impulses was greater than in the first event. As a result of the greater time between impulses, the event detection algorithm has split off the first impulse as a separate event when in fact it may be more appropriate for both events to be considered as a single event. Alternatively, it could be argued that each impulse should be considered a separate event and consequently the algorithm should have split event three into four separate events.

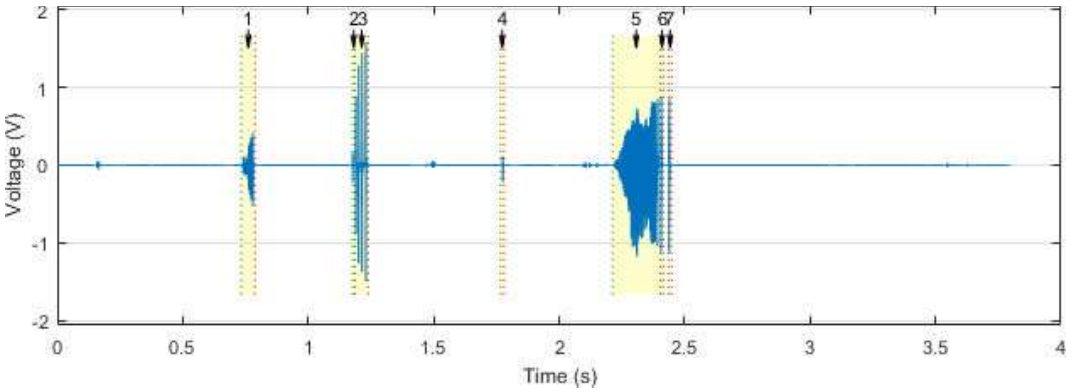

Figure 5.8: An AE recording from a second select participant (Participant 27, recording 2) with seven events, as identified by the event detection algorithm, shown bordered by dotted vertical lines and highlighted yellow.

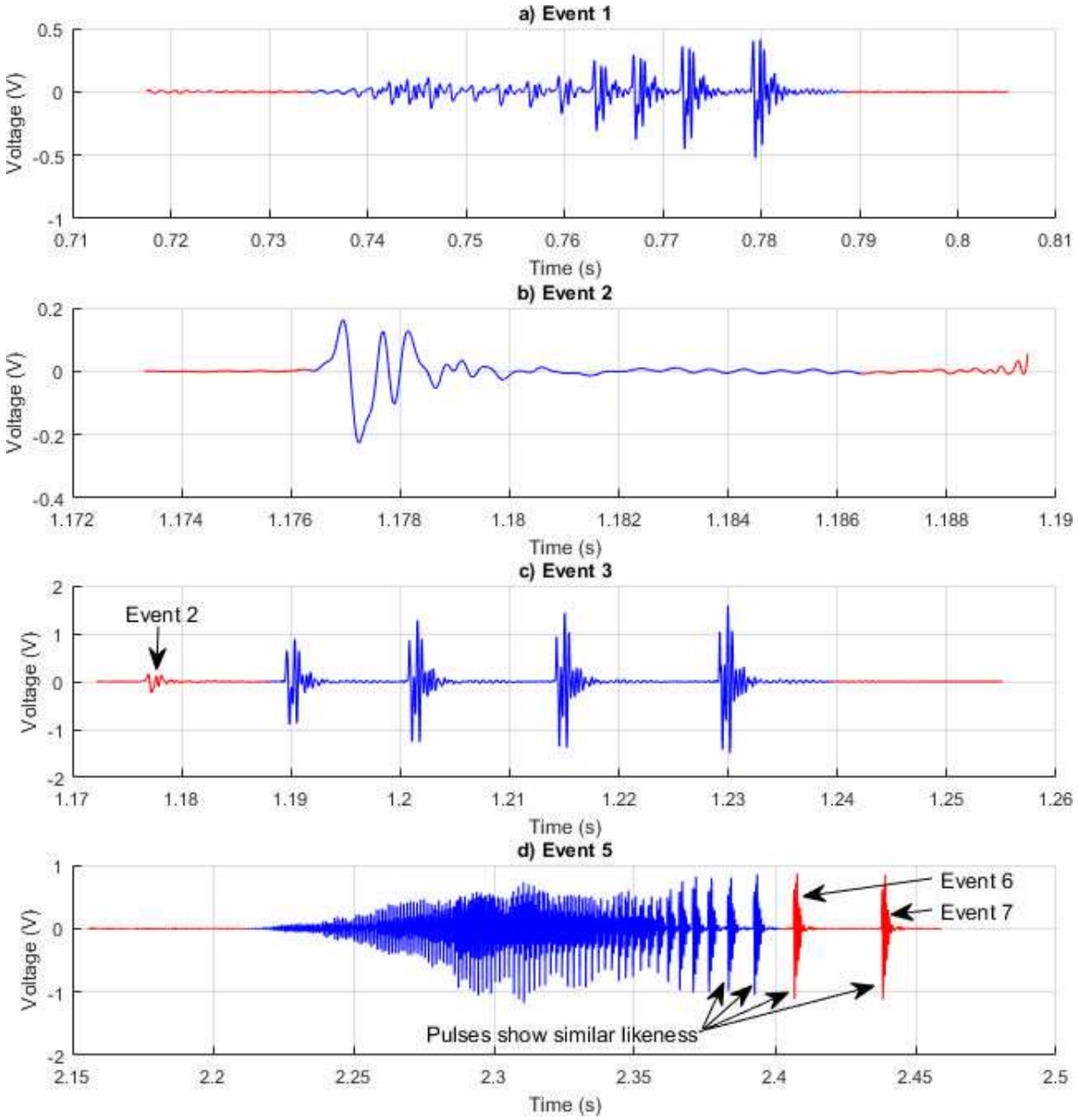

Figure 5.9: Event number (a) one, (b) two, (c) three, and (d) five from the signal of Figure 5.8. Note the event is coloured blue with the red signals being part of the original recording immediately surrounding each event.

Figure 5.9d shows the fifth event where it was observed that a similar sequence of successive decaying cycles to that of the second and third events has occurred. It was evident from observations of Figure 5.9d and the plots of the sixth and seventh events (not shown) that the two impulses that occurred immediately after the fifth event in Figure 5.9d (labelled Event 6 and Event 7 in the figure) were in fact identified as the sixth and seventh events by the event detection algorithm. Hence, the question was raised as to whether the sixth and seventh events should have been included into the fifth event due to their proximity and apparent likeness to

the pulses that occurred at the end of the fifth event (see annotation in Figure 5.9d identifying pulses with similar likeness). Conversely, it could be argued that the fifth event should have been split up further to separate more individual pulses from the end of the event as the signal voltage becomes very close to background noise voltages between the last few pulses of the fifth event. These results confirmed that when a series of successive events or pulses occur it will always be debatable as to where the cut-off between multiple events should lie or whether defining one long event would be more appropriate.

A range of different event detection threshold values were trialled and visual inspections of the event separation results were performed to determine the most suitable threshold values. For example, the event of Figure 5.9a contained a series of pulses and, after observing these pulses along with the signal surrounding them, it was determined that it was most appropriate to consider these pulses as a single event and so the algorithm threshold values were adjusted accordingly. However, in order for the events of Figure 5.9b and Figure 5.9c to become one event, the algorithm threshold values needed to be so low that merging of other arguably separate events occurred. The most suitable threshold values were determined to be those that provided a good balance between excluding long durations of low magnitude voltage values and avoiding extensive separation of pulse-like signals. As stated in Section 5.2, the value used for the initial variance threshold was 2.5 millivolts.

5.4 Event Metrics Summary

5.4.1 Introduction

A number of event signal metrics were computed in order to summarise each AE event by a limited number of representative metrics in place of a complete voltage-time signal. Due to the large bank of data and the 100 kHz recording rate, these representative metrics were needed to

reduce the data down to a manageable size and to make conclusions. The metrics computed were event duration, RMS amplitude, maximum amplitude, rise time, and main frequency content. Two additional metrics concerning the rate of occurrence of events, duration between consecutive events and the number of events per AE recording duration, were also determined. There were approximately 21,000 AE events detected by a sensor across the entire control and THR participant datasets with the median number of data points per event being approximately 650. By summarising each event by seven event metrics, there was an approximately 99% reduction in the amount of data representing the AE events. Consequently, while the data reduction does simplify subsequent analyses, it is important to keep in mind that such a significant data reduction has the potential to substantially limit the reliability of the subsequent analysis and conclusions that are drawn. Nevertheless, there still remained a large amount of data to analyse given the extensive number of AE events. The remainder of this section defines the specific definitions of the duration, peak amplitude, RMS amplitude, rise time, and main frequency content metrics as applied to the identified AE events.

The duration of an AE event signal was simply defined as the time elapsed in seconds between the first and last samples of the signal. For the AE event signals the sample rate was 100 kHz and therefore the time between samples was 1×10^{-5} seconds (0.01 milliseconds). Hence, the duration of an AE signal was given by Equation 5.1 as follows:

$$t_{dur} = (N - 1) \times (1 \times 10^{-5}) \tag{5.1}$$

Where, N was the number of samples in the signal and t_{dur} was the duration in seconds.

The peak amplitude of an AE event signal was defined as the maximum absolute voltage value of the signal. For the AE event signals each sample value was a voltage amplitude value in the range ±2.5 volts with a mean of zero volts. Hence, the voltage value furthest from zero in either

the positive or negative direction would be considered the peak amplitude. Therefore, the peak amplitude value was given by Equation 5.2 as follows:

$$A_{peak} = \max|x| \tag{5.2}$$

Where, x was the set of signal sample values and A_{peak} was the peak (maximum absolute) amplitude value. Additionally, the root-mean-squared (RMS) signal amplitude was defined by Equation 5.3 as follows:

$$A_{RMS} = \sqrt{\left[\sum_{i=1}^{N} x_i^2\right]/N} \tag{5.3}$$

Where, x_i was the voltage given by the i^{th} signal sample, N was the total number of samples in the signal, and A_{RMS} was the RMS voltage.

The rise time of an AE event signal was defined to be the duration in seconds from the beginning of the event to the occurrence of the peak amplitude of the event. This definition was used for computational simplicity as opposed to an alternate definition of, for example, the rise time being the duration between 10% and 90% of peak amplitude. For the AE event signals the rise time was calculated using Equation 5.4 as follows:

$$t_{rise} = (N_{peak} - 1) \times (1 \times 10^{-5}) \tag{5.4}$$

Where, N_{peak} was the number of samples from the beginning of the signal to the peak amplitude sample, the constant 1×10^{-5} was the time between samples in seconds, and t_{rise} was the rise time in seconds.

As mentioned in the previous chapter, the main frequency of a signal was defined as the frequency at which the magnitude of the Fourier transform of the signal was greatest. For each AE event, the Fourier transform was applied using the Fast Fourier Transform (FFT) algorithm to yield the frequency domain response of the event. The frequency value that corresponded to the maximum frequency domain magnitude was then taken as the main frequency content metric.

5.4.2 Empirical Cumulative Distribution & Probability Densities

In an effort to identify and investigate trends in the AE data, the computed metrics were first grouped by the revision reasons, then empirical cumulative distribution functions (CDFs) of each grouped metric were calculated, and finally the data plotted for visual comparisons. Figure 5.10 presents the empirical CDFs of the control participants and THR participants with the noise, loosening, and wear revision reasons for each of the signal metrics. There were a number of important observations from the plots of Figure 5.10. The most noticeable feature observed from the empirical CDFs was that the metric values spanned several orders of magnitude, for all but one metric. Therefore, the metric value axes of the empirical CDFs, excluding that of the number of events per recording duration CDF, were presented using a log scale to more clearly show distinctions between the CDFs of the different groups. Conversely, while the metric values overall did span several orders of magnitude, it was typically observed from the empirical CDFs that a substantial proportion of the events for each metric had values within a relatively narrow range, as indicated in Figure 5.10 by a narrow region of metric values with a steep gradient for the first five of the seven metrics.

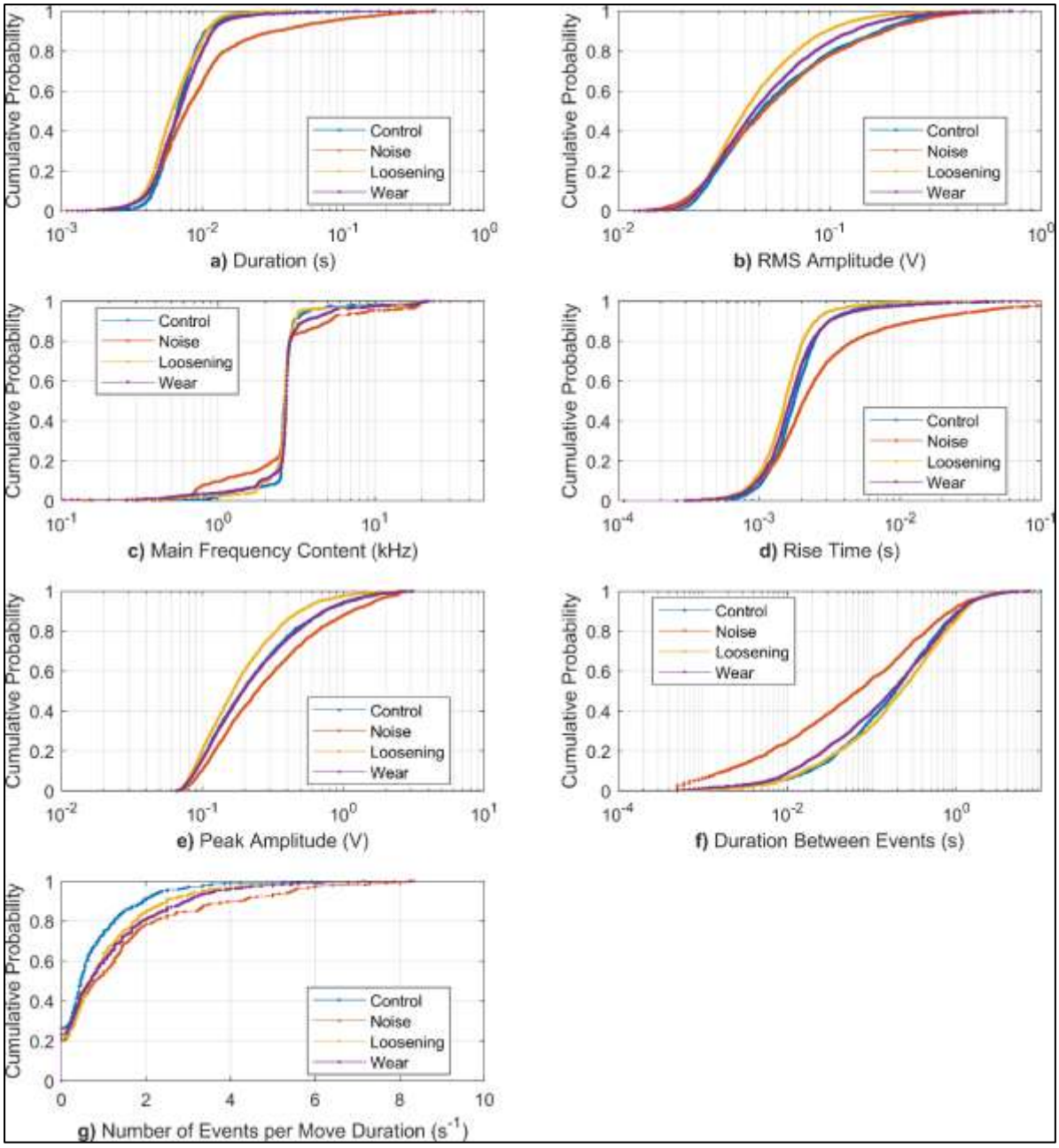

Figure 5.10: Empirical cumulative distribution functions (CDFs) of acoustic event metric a) duration, b) RMS amplitude, c) main frequency content, d) rise time, e) peak amplitude, f) duration between events, and g) number of events per recording duration. Each subfigure contains a CDF for the control group and the noise, loosening, and excessive wear revision reasons.

To more thoroughly investigate the distribution of the AE metrics, probability density histograms were computed of the AE metrics. These histograms used the same groupings presented in the empirical CDFs of Figure 5.10 (i.e. control participants, noise revision, loosening revision, and excessive wear revision reasons) as well as an additional group containing all THR participant AE events. The probability density histograms allow a more natural interpretation of the distributions of the datasets and place an emphasis on comparisons

of the main representative portions of the data, whereas the empirical CDFs emphasised differences in the extremities of the datasets. That is to say, the histograms are dominated by data around the central tendency and the data at the extremities is much less prominent and sometimes unnoticed.

The histograms were normalised to make the total area of each histogram equal to one (i.e. integration of these histograms would produce the corresponding CDFs). Additionally, for each metric in turn, the histograms across each different grouping could be directly compared since the same bin size and locations were used for the computations. The histograms for the event duration, RMS amplitude, main frequency content, and number of events per recording duration are shown in Figures 5.11-5.14 respectively. Furthermore, as a quantitative summary of the distributions, Tables 5.1 and 5.2 present the 1^{st}, 25^{th}, 50^{th}, 75^{th}, and 99^{th} percentile values, the mean value, and the interquartile range (i.e. 75^{th} minus 25^{th} percentile) for the AE data of each group and for each of the metrics. By observing the histograms and summary percentiles, in addition to the empirical CDFs, the distributions of the AE event data metrics were more apparent and a better understanding of the general nature of the AE events for each grouping was gained through quantitative comparisons. Note that the histograms of the rise times, peak amplitudes, and duration between events have not been presented here as they did not offer any substantial further insight in addition to what was gained from the other four metrics.

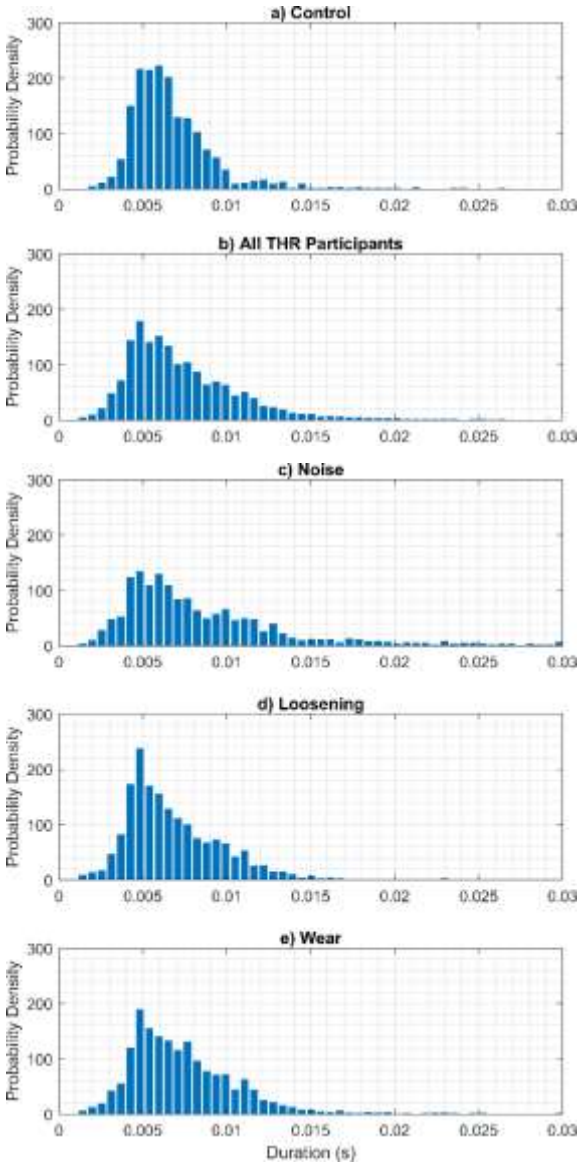

Figure 5.11: Histogram of AE event durations for a) control, b) all THR, c) noise revision reason, d) loosening revision reason, and e) excessive wear revision reason participants.

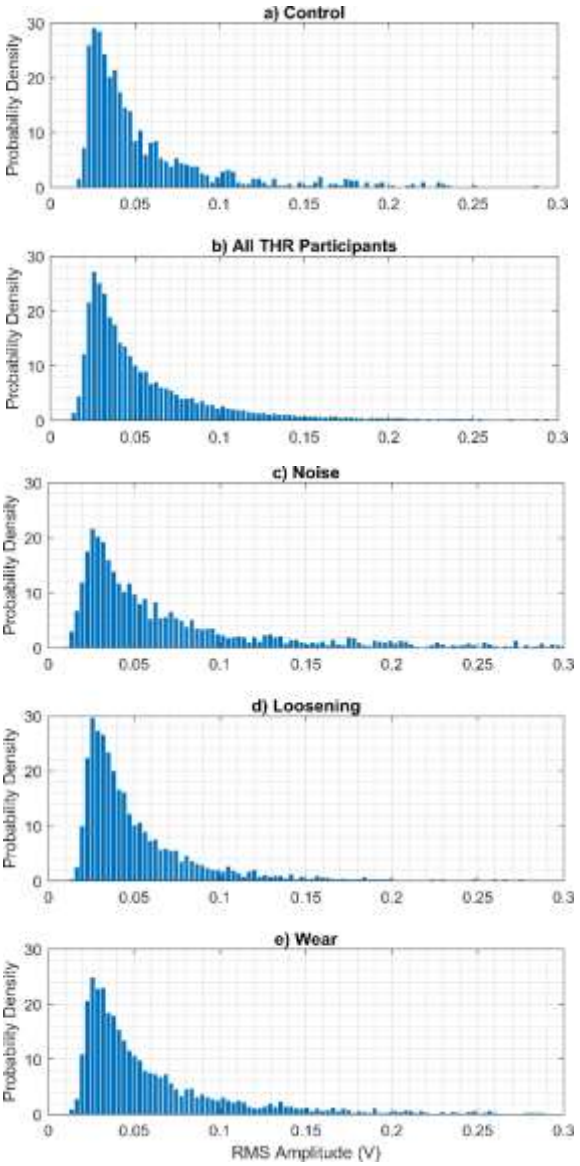

Figure 5.12: Histogram of AE event root-mean-squared amplitudes for a) control, b) all THR, c) noise revision reason, d) loosening revision reason, and e) excessive wear revision reason participants.

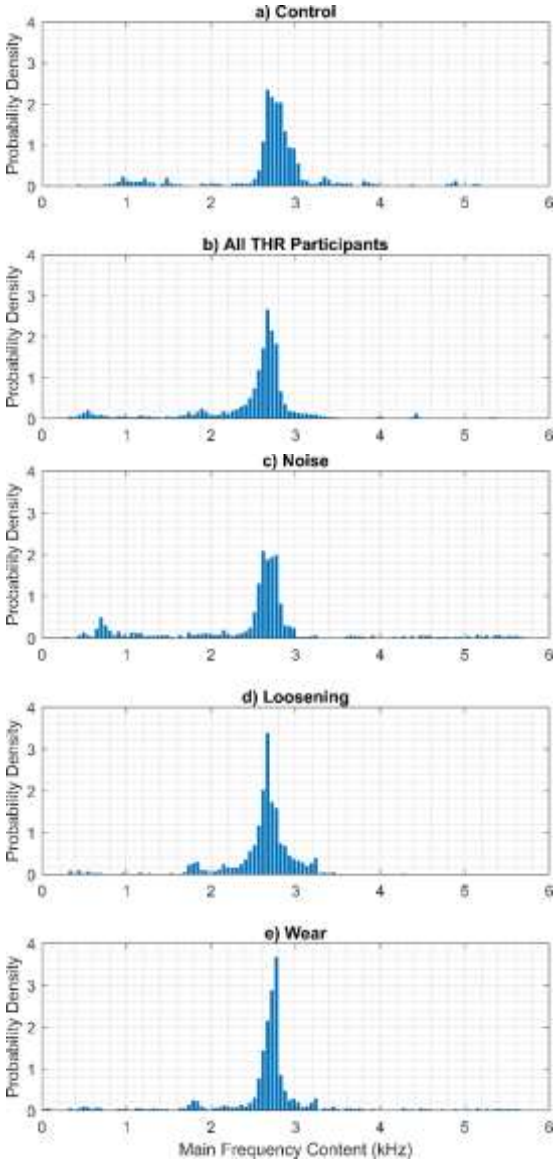

Figure 5.13: Histogram of AE event main frequency content for a) control, b) all THR, c) noise revision reason, d) loosening revision reason, and e) excessive wear revision reason participants.

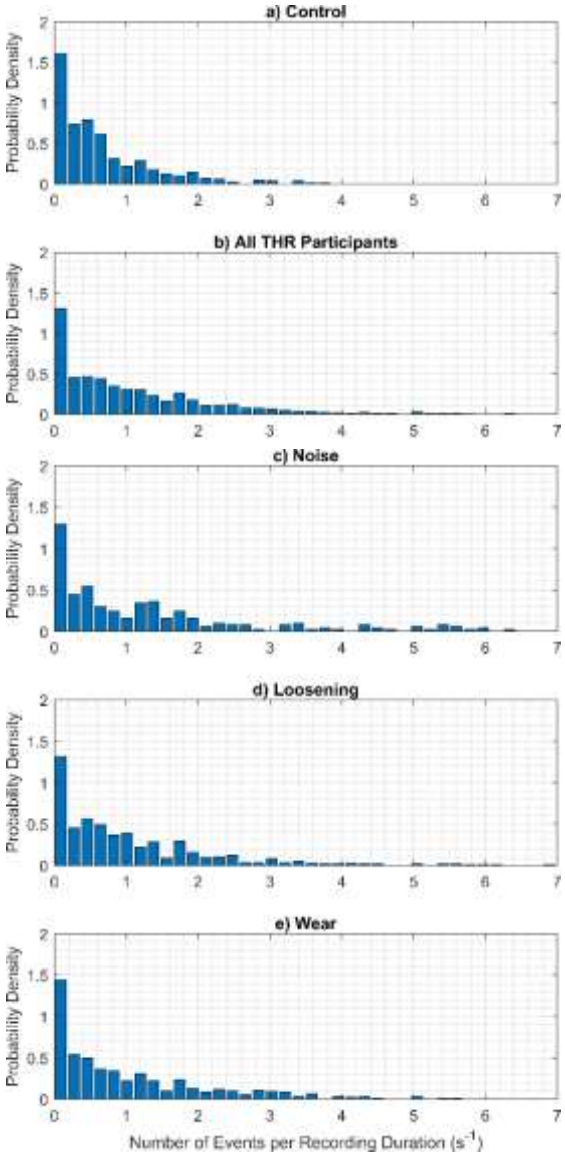

Figure 5.14: Histogram of number of AE events per recording duration for a) control, b) all THR, c) noise revision reason, d) loosening revision reason, and e) excessive wear revision reason participants.

Table 5.1: Percentile, interquartile range, and mean values of four event metrics for the control, noise revision, loosening revision, excessive wear revision groups, and all THR participants.

	Group	Percentile					Interquartile Range	Mean
		1st	25th	50th (Median)	75th	99th		
Event Duration (ms)	Control	2.9	5.1	6.2	7.8	24.3	2.7	7.2
	Noise	2.3	5.3	7.7	12.3	266.9	7.0	20.0
	Loosening	2.1	4.8	6.1	8.5	18.4	3.7	7.0
	Excessive Wear	2.2	5.1	6.9	9.3	40.4	4.2	8.9
	All THR Participants	2.3	5.0	6.7	9.5	123.2	4.5	11.0
Event RMS Amplitude (mV)	Control	19.7	29.2	40.5	64.5	478.8	35.3	61.3
	Noise	15.5	30.5	48.9	90.6	442.6	60.1	80.8
	Loosening	18.5	29.3	39.9	61.8	229.9	32.5	53.6
	Excessive Wear	17.8	30.2	44.1	73.8	341.3	43.6	65
	All THR Participants	16.8	29.2	42.2	71.0	384.4	41.8	65.0
Event Main Frequency Content (kHz)	Control	0.75	2.66	2.77	2.92	19.98	0.26	3.14
	Noise	0.49	2.52	2.67	2.81	19.42	0.29	3.41
	Loosening	0.44	2.55	2.67	2.80	19.02	0.25	2.99
	Excessive Wear	0.35	2.61	2.73	2.80	19.20	0.19	3.15
	All THR Participants	0.41	2.47	2.66	2.77	20.45	0.30	3.16
Number of Events per Recording Duration (s^{-1})	Control	0	0	0.45	0.94	3.33	0.94	0.65
	Noise	0	0.21	0.79	1.79	7.94	1.58	1.44
	Loosening	0	0.21	0.71	1.46	5.92	1.25	1.05
	Excessive Wear	0	0.14	0.68	1.67	6.12	1.53	1.13
	All THR Participants	0	0.21	0.77	1.67	6.57	1.46	1.17

Table 5.2: Percentile, interquartile range, and mean values of three event metrics for the control, noise revision, loosening revision, excessive wear revision groups, and all THR participants.

	Group	Percentile					Interquartile Range	Mean
		1st	25th	50th (Median)	75th	99th		
Event Rise Time (ms)	Control	0.6	1.3	1.7	2.2	11.8	0.9	2.2
	Noise	0.5	1.4	2.1	3.7	142.1	2.3	8.4
	Loosening	0.6	1.2	1.5	1.9	8.2	0.7	1.8
	Excessive Wear	0.5	1.3	1.7	2.2	21.0	0.9	2.5
	All THR Participants	0.6	1.3	1.8	2.4	62.6	1.1	3.9
Event Peak Amplitude (mV)	Control	70	110	164	288	2424	178	276
	Noise	74	137	245	529	2527	392	459
	Loosening	71	107	165	284	1478	177	251
	Excessive Wear	71	118	195	388	2101	270	335
	All THR Participants	71	115	185	359	2447	244	327
Duration Between Events (ms)	Control	1.4	95.9	331.1	748.0	3963	652.1	585.9
	Noise	0.5	10.6	69.4	313.4	2873	302.8	306.4
	Loosening	1.7	60.1	224.4	617.2	3119	557.1	461.8
	Excessive Wear	0.8	39.0	171.3	548.7	3421	509.7	426.6
	All THR Participants	0.6	44.0	191.2	562.2	3253	518.2	433.7

From observing the empirical CDFs, probability density histograms, and the summary statistics of Tables 5.1 and 5.2 a positive skew of all the data across all metrics and for all groupings was apparent. The positive skew in the data was obvious from observing that, for all metrics and all groups, the mean values were greater than the median values. Additionally, the relative skewness between groups (indicated by a larger difference in the mean and median values) was increased for the groups with increased interquartile ranges. The positive skew of the data is important to note as it indicates that the data is not normally distributed. For all the metrics of Tables 5.1 and 5.2 the data can be skewed upwards. However, the data cannot be skewed very far downwards since negative values make no physical sense for the metrics. The skewed

characteristics of the metrics initially suggest a lognormal distribution may be present. While the lognormal distribution did appear to fit the data well for some metrics of some of the participant groups, the lognormal distribution was not always a good fit for most of the data. Consequently, treating any metrics as lognormally distributed would not necessarily aid analysis and care needs to be taken when applying any statistical analysis techniques to the metrics.

Nevertheless, presenting the empirical CDFs with a log scale further distinguished another noticeable feature of the plots. Figure 5.10 clearly shows the empirical CDF of the noise revision reason group, across all metrics, was noticeably different than the empirical CDFs of the other groups. This observation can largely be explained by considering that the noise revision reason group contained the participants whose implants produced audible squeaking. The audible squeaking events had considerably longer durations than the majority of other events that were detected as identified in Section 4.3. Figure 5.10a confirms the occurrence of longer durations by showing that the gradient of the noise group CDF began to substantially decrease at a lower duration than the other groups and hence indicated there was a higher proportion of events from the noise group that had longer durations. In the event duration histograms of Figure 5.11 it was observed that the noise revision group had a lower magnitude and more broad overall probability density peak than the other groups, especially compared to the control group, indicating a greater spread of event durations in the majority of the noise revision group data. The greater spread of the noise revision group was quantified by the results in Table 5.1 where the duration interquartile ranges (IQRs) were 7.0 ms, 3.7 ms, and 2.7 ms for the noise revision, loosening revision, and the control groups respectively.

Additionally, given that for an event to be audible it must exceed a certain energy threshold and have sufficient duration for it to be perceived, the increased durations at sustained increased

amplitude levels of the audible squeaking events led to increased RMS voltage amplitudes. Figure 5.10b confirms that the noise revision reason group had more occurrences of events with increased RMS voltage amplitudes. The presence of increased RMS voltage amplitudes of the noise group was more subtle in the histograms of event RMS voltage amplitudes in Figure 5.12. However, the peak of the probability density histogram for the noise group in Figure 5.12 was noticeably lower than the other groups which indicated greater probability density magnitudes for larger RMS voltage amplitudes. Table 5.1 quantifies the increased RMS voltage amplitudes of the noise revision group by showing median values of 48.9 mV, 40.5 mV, and 39.9 mV for the noise revision, control, and loosening revision groups respectively. Furthermore, Table 5.1 shows an interquartile range of the noise group that was approximately 70% larger than that of the control group.

The rise time empirical CDF (Figure 5.10d) of the noise revision reason group indicated substantially more events with larger rise times than those of the other groups. It was observed from initial observations that, due to the way the rise time was defined, rise times for squeaking events would often equate to a considerable portion of the duration of the event. Hence, given there was a larger proportion of noise group events with increased durations, it follows that there would be a corresponding larger proportion of noise group events with longer rise times. Table 5.2 demonstrates the increased rise times of the noise group quantitatively by showing median rise times of 2.1 ms, 1.7 ms, and 1.5 ms for the noise revision, control, and loosening revision groups respectively.

The main frequency content CDFs of Figure 5.10c showed broadly similar frequency behaviour across all groups. With respect to main frequency content, all of the groups had at least the middle 50% of their events within the range of approximately 2.5-3.0 kHz. However, despite the similar central tendency values of the different groups, the noise revision reason group was

observed to have a wider spread of event main frequency content away from the median value (the median value was 2.67 kHz) in both directions. This wider spread of main frequency content was not immediately obvious in the histograms of Figure 5.13 as the largest probability density magnitudes were observed in a very similar range of frequencies. Furthermore, the summary statistics of Table 5.1 did not emphasise the wider spread of the main frequency content of the noise revision group. It was observed from Table 5.1 that the main frequency content of the noise group had frequencies of 2.52 kHz, 2.67 kHz, and 2.81 kHz for the 25^{th}, 50^{th}, and 75^{th} percentiles respectively and all the other groups had values that were within 150 Hz of the aforementioned frequencies for the same percentiles.

The empirical CDF of the duration between events (Figure 5.10f) showed that for the noise revision reason group there was an obvious greater occurrence of shorter durations between consecutive events. It was also observed that the control group had more occurrences of longer durations between events which distinguished it from the remaining revision reason groups of loosening and excessive wear. These observations of the difference between groups in durations between consecutive events were quantitatively reflected in Table 5.2. The median duration between consecutive events of the control group was 331.1 ms which was over 4.5 times longer than the median duration between consecutive events of the noise revision group (69.4 ms) and approximately 1.5 times longer than the median duration between consecutive events of the loosening revision group (224.4 ms).

The main observation from the empirical CDFs of the number of events per recording duration (Figure 5.10g) was that the control group showed a distinct difference to the revision reason groups. The control group showed a greater proportion of data with a lower number of events per recording duration than the revision reason groups. This observation verifies the initial observations presented in Section 4.3, where it was observed that the amount of acoustic

activity was noticeably less for the control AE recordings than for the THR participant AE recordings. The histogram of Figure 5.14 confirms the observation of the control group having a greater proportion of data with a lower number of events per recording duration than the revision reason groups by showing the control group to have noticeably greater probability density magnitudes for the four lowest histogram bins than the corresponding histogram bins of the other groups. However, the histograms did not convey well an important observation for this metric that was evident in the empirical CDFs and Table 5.1; at least 20% of the AE recordings for all groups had zero events detected by the event detection algorithm. It is important to note that variation in the recording durations of the AE recordings was present due to the self-selected movement speeds of the motions performed by the participants during testing. This variation in recording durations should be kept in mind when making conclusions regarding the metrics presented here since the occurrence of AE events may be correlated to the specific movement of the participant's lower limb irrespective of the time taken to make the movement.

In general, the empirical CDF plots and probability density histograms suggest that THR participants that are categorised into the noise revision reason group have empirical CDFs and probability densities for most metrics that are noticeably different from those of the other groups. The empirical CDFs and probability density histograms for the main frequency content of events seemed to be generally similar for all groups. However, for the noise revision group there was a noticeable difference in the size of the tails of the main frequency content distribution indicating a larger spread of the data. Additionally, it was observed from the data of duration between events and number of events per recording duration that AE events occurred more frequently in the data from the THR participant groups than the control group.

5.4.3 Kolmogorov-Smirnov Tests – Control and Revision Reasons

In order to more thoroughly confirm any differences in the AE event data from the control and different revision reason groups a statistical test was employed. A two-sample Kolmogorov-Smirnov (KS) test was applied to the data for each of the metrics presented in Figure 5.10. The two-sample KS test is a nonparametric statistical test which evaluates the null hypothesis that the empirical CDFs of two datasets come from the same distribution. The KS test does this comparison by first evaluating what is known as the test statistic: the maximum absolute difference between the CDFs of the two distributions. The KS test then calculates the probability (*p* value) of observing a test statistic that is as extreme as, or more extreme than, the test statistic observed in the first step. The KS test will then reject the null hypothesis if the calculated *p* value is less than the specified significance level, α, where α is commonly either 0.05 (5%) or 0.01 (1%).

As explained above, the two-sample KS test compares only two CDFs at one time. Therefore, to apply the KS test to all the data of Figure 5.10 the test needed to be applied multiple times. Consequently, for each metric, six different KS tests were conducted to cover the six possible different combinations of pairs for the four groups. Tables 5.3-5.6 present the results of the KS tests for each event metric and using a significance level of 1%. It was observed from the KS test results that in all but 6 of the 42 tests the null hypothesis (that the tested datasets belonged to the same distribution) was rejected. A rejected null hypothesis indicates that enough evidence was present to suggest the null hypothesis was false and the alternative hypothesis (that the datasets tested come from different distributions) can be accepted.

Table 5.3: Kolmogorov-Smirnov test results for event duration and RMS amplitudes of control and revision reason group participants. A significance level of 1% was used.

Pairing	Duration			RMS Amplitude		
	Reject Null Hypothesis	P value	Test Statistic	Reject Null Hypothesis	P value	Test Statistic
Control-Noise:	1	<0.001	0.286	1	<0.001	0.131
Control-Loosening:	1	<0.001	0.084	0	0.421	0.032
Control-Wear:	1	<0.001	0.142	1	0.002	0.066
Noise-Loosening:	1	<0.001	0.215	1	<0.001	0.144
Noise-Wear:	1	<0.001	0.169	1	<0.001	0.077
Loosening-Wear:	1	<0.001	0.096	1	<0.001	0.075

Table 5.4: Kolmogorov-Smirnov test results for event main frequency content and rise time of control and revision reason group participants. A significance level of 1% was used.

Pairing	Main Frequency Content			Rise Time		
	Reject Null Hypothesis	P value	Test Statistic	Reject Null Hypothesis	P value	Test Statistic
Control-Noise:	1	<0.001	0.224	1	<0.001	0.237
Control-Loosening:	1	<0.001	0.241	1	<0.001	0.133
Control-Wear:	1	<0.001	0.187	0	0.303	0.034
Noise-Loosening:	1	<0.001	0.113	1	<0.001	0.313
Noise-Wear:	1	<0.001	0.156	1	<0.001	0.223
Loosening-Wear:	1	<0.001	0.179	1	<0.001	0.104

Table 5.5: Kolmogorov-Smirnov test results for event peak amplitude and duration between events of control and revision reason group participants. A significance level of 1% was used.

Pairing	Peak Amplitude			Duration Between Events		
	Reject Null Hypothesis	P value	Test Statistic	Reject Null Hypothesis	P value	Test Statistic
Control-Noise:	1	<0.001	0.196	1	<0.001	0.315
Control-Loosening:	0	0.816	0.023	1	<0.001	0.106
Control-Wear:	1	<0.001	0.099	1	<0.001	0.156
Noise-Loosening:	1	<0.001	0.195	1	<0.001	0.237
Noise-Wear:	1	<0.001	0.104	1	<0.001	0.175
Loosening-Wear:	1	<0.001	0.100	1	<0.001	0.074

Table 5.6: Kolmogorov-Smirnov test results for number of events per recording duration of control and revision reason group participants. A significance level of 1% was used.

Pairing	Number of Events per Recording Duration		
	Reject Null Hypothesis	P value	Test Statistic
Control-Noise:	1	<0.001	0.245
Control-Loosening:	1	<0.001	0.173
Control-Wear:	1	<0.001	0.175
Noise-Loosening:	0	0.019	0.111
Noise-Wear:	0	0.175	0.079
Loosening-Wear:	0	0.330	0.052

For the KS tests of the metrics of event duration (Table 5.3), event main frequency content (Table 5.4), and duration between events (Table 5.5), it was observed that the null hypothesis was rejected for all tests. Hence, the KS test concluded with a 99% confidence that, for each of the three aforementioned metrics, the difference in the datasets of the control and revision

reason groups was statistically significant and consequently it could be concluded they all came from separate distributions. For the KS tests of the metrics of event RMS (Table 5.3) and peak (Table 5.5) amplitudes it was observed that the null hypothesis was rejected for all group parings except for the control and loosening group pair. For the KS tests of the event rise time metric (Table 5.4) it was observed that there was a failure to reject the null hypothesis for the control and excessive wear paring only. For the KS tests of the number of events per recording duration metric (Table 5.6) it was observed that there was a failure to reject the null hypothesis for the noise and loosening, the noise and excessive wear, and the loosening and excessive wear pairings. The failure to reject the null hypothesis for these six group pairing and metric combinations indicates there was not enough evidence available to conclude that the datasets concerned came from distributions that had a statistically significant difference.

The KS tests indicated statistically significant differences in most of the distributions of the datasets from the control and THR revision reason groups. This finding was important as it made progress in answering two of the research questions posed in Section 1.5; could AEs from THR implants be reliably categorised using time and frequency domain characteristics, and would those categories corresponded to the clinical revision reasons? The results from the KS tests suggested that failing THR implants produced some AEs that were unique to the implant's particular failure mode. Hence, the KS test results did suggest that THR AEs could be categorised into clinical revision reason groups using time and frequency domain characteristics. However, the results presented thus far do not indicate whether the categorisation of AEs can be performed reliably. To answer the research questions more rigorously, and validate the reliability of the categorisation of AEs, the distribution of the AEs of the individual group participants were compared to the overall distribution of the group to which they were categorised. That is, using the KS test, the empirical CDFs of the individual

participants of a group were compared against the empirical CDF of the entire group to check for any statistically significant differences.

5.4.4 Kolmogorov-Smirnov Tests – Individual Control

Tables 5.7-5.10 show the KS test results for the seven metrics using the pairings of the overall empirical CDF of the control group with the empirical CDF of each control participant individually. A significance level of 1% was used for all KS tests of Tables 5.7 -5.10. Given that it was assumed that the data from each of the control participants would be similar (thus characteristic of AEs not produced by a THR implant), it would be expected that the KS tests would fail to reject the null hypothesis and not find statistically significant differences. However, it should be kept in mind that failing to reject the null hypothesis only indicates that not enough evidence was present to reject the null hypothesis and does not confirm the tested distributions are from the same distribution. Nevertheless, to support the assumption that the control data was well categorised, the data of Tables 5.7 -5.10 would not be expected to show many, if any, statistically significant differences in the results.

Table 5.7: Kolmogorov-Smirnov test results for event duration and RMS amplitudes of control group participants. A significance level of 1% was used.

Pairing	Duration			RMS Amplitude		
	Reject Null Hypothesis	P value	Test Statistic	Reject Null Hypothesis	P value	Test Statistic
Control-C1:	0	0.073	0.244	0	0.818	0.120
Control-C2:	0	0.643	0.113	0	0.825	0.096
Control-C3:	0	0.038	0.288	0	0.028	0.299
Control-C4:	0	0.168	0.398	0	0.403	0.319
Control-C5:	0	0.265	0.251	0	0.022	0.376
Control-C6:	0	0.262	0.174	0	0.308	0.167
Control-C7:	0	0.625	0.095	0	0.790	0.082
Control-C8:	0	0.171	0.131	0	0.063	0.156
Control-C9:	0	0.464	0.133	0	0.606	0.119
Control-C10:	0	0.255	0.389	0	0.369	0.352
Control-C11:	0	0.582	0.235	0	0.640	0.225
Control-C12:	0	0.075	0.431	0	0.057	0.448
Control-C13:	0	0.120	0.424	0	0.462	0.304
Control-C14:	0	0.233	0.131	0	0.013	0.201
Control-C15:	1	<0.001	0.545	1	<0.001	0.477
Control-C16:	0	0.035	0.148	0	0.108	0.126
Control-C17:	0	0.692	0.067	0	0.024	0.139
Control-C18:	0	0.117	0.209	0	0.306	0.170
Control-C19:	1	<0.001	0.181	0	0.610	0.068
Control-C20:	0	0.298	0.164	0	0.301	0.163
Control-C21:	0	0.245	0.086	1	0.006	0.143
Control-C22:	1	0.005	0.308	0	0.106	0.215

Table 5.8: Kolmogorov-Smirnov test results for event main frequency content and duration between events of control group participants. A significance level of 1% was used.

Pairing	Main Frequency Content			Duration between Events		
	Reject Null Hypothesis	P value	Test Statistic	Reject Null Hypothesis	P value	Test Statistic
Control-C1:	0	0.283	0.187	0	0.081	0.340
Control-C2:	1	0.008	0.254	0	0.458	0.160
Control-C3:	1	<0.001	0.649	0	0.126	0.316
Control-C4:	1	0.007	0.604	0	0.205	0.663
Control-C5:	1	<0.001	0.549	0	0.031	0.517
Control-C6:	1	<0.001	0.493	0	0.266	0.231
Control-C7:	1	<0.001	0.527	1	<0.001	0.359
Control-C8:	1	<0.001	0.377	1	<0.001	0.287
Control-C9:	1	<0.001	0.445	0	0.733	0.141
Control-C10:	1	<0.001	0.794	0	0.539	0.499
Control-C11:	1	0.009	0.500	0	0.395	0.558
Control-C12:	1	<0.001	0.760	0	0.298	0.510
Control-C13:	1	0.008	0.596	0	0.099	0.997
Control-C14:	1	<0.001	0.256	0	0.916	0.088
Control-C15:	1	<0.001	0.452	1	<0.001	0.295
Control-C16:	1	<0.001	0.322	0	0.188	0.125
Control-C17:	1	<0.001	0.433	1	0.007	0.174
Control-C18:	1	<0.001	0.533	0	0.144	0.297
Control-C19:	1	0.001	0.174	0	0.158	0.111
Control-C20:	1	<0.001	0.385	1	0.008	0.419
Control-C21:	1	<0.001	0.314	0	0.270	0.092
Control-C22:	1	<0.001	0.387	0	0.383	0.243

Table 5.9: Kolmogorov-Smirnov test results for event rise time and peak amplitude of control group participants. A significance level of 1% was used.

Pairing	Rise Time			Peak Amplitude		
	Reject Null Hypothesis	P value	Test Statistic	Reject Null Hypothesis	P value	Test Statistic
Control-C1:	0	0.701	0.134	0	0.201	0.203
Control-C2:	0	0.166	0.170	0	0.391	0.137
Control-C3:	0	0.362	0.189	0	0.019	0.312
Control-C4:	0	0.900	0.204	0	0.838	0.221
Control-C5:	0	0.030	0.362	0	0.096	0.309
Control-C6:	0	0.158	0.194	0	0.535	0.139
Control-C7:	1	0.002	0.234	0	0.176	0.140
Control-C8:	1	<0.001	0.282	0	0.030	0.172
Control-C9:	0	0.041	0.217	0	0.253	0.158
Control-C10:	0	0.738	0.262	0	0.558	0.304
Control-C11:	0	0.675	0.219	0	0.368	0.278
Control-C12:	0	0.874	0.199	0	0.125	0.395
Control-C13:	0	0.328	0.339	0	0.271	0.357
Control-C14:	0	0.312	0.122	1	0.002	0.232
Control-C15:	1	<0.001	0.675	1	<0.001	0.319
Control-C16:	1	0.006	0.178	0	0.333	0.098
Control-C17:	0	0.058	0.124	0	0.244	0.096
Control-C18:	0	0.015	0.273	0	0.147	0.200
Control-C19:	0	0.071	0.116	0	0.367	0.082
Control-C20:	0	0.808	0.107	0	0.284	0.166
Control-C21:	1	0.008	0.139	0	0.018	0.128
Control-C22:	0	0.705	0.125	1	0.004	0.312

Table 5.10: Kolmogorov-Smirnov test results for number of events per recording duration of control group participants. A significance level of 1% was used.

Pairing	Number of Events per Recording Duration		
	Reject Null Hypothesis	P value	Test Statistic
Control-C1:	0	0.066	0.330
Control-C2:	0	0.149	0.272
Control-C3:	0	0.076	0.323
Control-C4:	1	<0.001	0.539
Control-C5:	0	0.046	0.328
Control-C6:	0	0.736	0.163
Control-C7:	0	0.060	0.260
Control-C8:	1	0.002	0.390
Control-C9:	0	0.390	0.184
Control-C10:	1	<0.001	0.567
Control-C11:	1	0.007	0.414
Control-C12:	1	<0.001	0.515
Control-C13:	1	<0.001	0.461
Control-C14:	0	0.199	0.207
Control-C15:	1	<0.001	0.427
Control-C16:	0	0.037	0.321
Control-C17:	1	<0.001	0.663
Control-C18:	0	0.225	0.221
Control-C19:	1	<0.001	0.622
Control-C20:	0	0.332	0.186
Control-C21:	1	<0.001	0.655
Control-C22:	0	0.229	0.208

It was observed from Table 5.7 that for the event RMS amplitude KS tests there was a statistically significant difference in the data for only two of the 22 participants. For the event duration and peak amplitude KS tests (Tables 5.7 and 5.9 respectively), there was a statistically significant difference in the data for only three participants for each metric. For the KS tests of duration between events (Table 5.8) and those of the event rise time (Table 5.9) five participants had data with a statistically significant difference. The KS tests for the number of events per recording duration (Table 5.10) had 10 participants with a statistically significant difference in their data. Furthermore, the main frequency content KS tests of Table 5.8 had 21 of the 22 datasets with a statistically significant difference.

It would seem that, from the KS test results of the control group data, a number of the control participants have AE data that does not fit well with the group as a whole. However, when considering the recruitment criteria for the control group it would be unwise to remove participants' AE data from the control group based solely on the results from the KS tests. The recruitment criteria for the control group required participants to have healthy natural hip joints with no history of hip joint trauma or injury. Hence, to get a comprehensive indication of the AEs from natural hips, data should only be removed from the control group if it was erroneous and not simply because it had AEs with differing metrics to the other control participants' data.

Nevertheless, performing the KS tests on the control data was not meaningless. From the data of Tables 5.7-5.10, control participant 15 was noted to be the only participant to have a statistically significant difference for all of the metrics. The empirical CDFs were then investigated and it was found that control participant 15 exhibited behaviour that was noticeably different to that of the other control participants. Upon further inspection of the time domain responses of control participant 15's AEs it was found that there was an obvious increase in noise present. This increased noise was likely due to inadequate force holding the

AE sensor pad in place during testing causing AEs to be detected from the movement of the sensor pad. For these reasons, it was decided that the data of control participant 15 would be removed from all further analysis.

A further important observation from the KS tests of the control data was the result that 21 of the 22 main frequency content datasets showed a statistically significant difference and 16 of these tests yielded *p* values below 0.001 (see Table 5.8). This result was initially unexpected as none of the other metrics showed such a trend and, as identified in the initial observations of Chapter 4, the frequency responses were shown to primarily cover a relatively narrow frequency range. Conversely, it was also identified in the observations of Chapter 4 that the main frequency content did vary (sometimes considerably) between participants and hence some variability in their distributions was expected. Therefore, the result of the KS tests in Table 5.8 emphasise the variability and overall data spread that was present in the main frequency content of the AEs was in fact significant. Thus, it was noted for future analysis that significant variation in the main frequency content was expected and that any comparisons of the frequency content should be carried out with this in mind.

5.4.5 Kolmogorov-Smirnov Tests – Individual Noise Revisions

Further KS tests were performed for the seven metrics using the pairings of the overall empirical CDF of the noise revision reason group with the empirical CDF of each noise revision reason participant individually. Additionally, KS tests were performed between the individual noise revision reason participants and the overall control group to investigate similarities of the individual noise revision participants to the control group. There were nine participants who had been categorised into the noise revision group. Hence, two sets of nine KS tests were carried out to compare each individual noise revision group participant to the overall noise

revision group and control group empirical CDFs. The results of these KS tests are shown in Table B.1 to Table B.4 in Appendix B and a significance level of 1% was used for all the tests.

The results from both sets of KS tests showed that a statistically significant difference was present for all but one of the 18 tests of the main frequency content event metric. This result was consistent with the previous finding of statistically significant differences in the main frequency content of the individual control group participants. Across the six other metrics, for the set of KS tests comparing the noise revision group to its individual participants, an average of six out of the nine tests showed statistically significant differences. This result is important as it showed that, in general, for each individual metric over half of the participants categorised into the noise revision group showed evidence that they had AE behaviour that differed significantly from the overall behaviour of the group itself. Hence, the result implies that there was significant variation in the AEs of participants who were identified as being in the noise revision group. This is expected due to the range of possible sources of noise that may come from an implant.

Furthermore, for the set of KS tests comparing the control group to the individual noise revision group participants, an average of five tests per metric (excluding main frequency content) showed statistically significant differences. This result implies that for each metric (excluding main frequency content) the AE behaviour of just under half of the noise revision group participants was not confirmed to be statistically different to the behaviour of control group AEs. Hence, more analysis would need to be carried out to investigate the possibility of noise group participants having AE content that was similar to that of the control group AEs.

Interestingly, two of the noise group participants (Participants 10 and 25) were found to have statistically significant differences to the overall noise group in all metrics except RMS

amplitude. When comparing the KS tests using the control group data pairing (and excluding the main frequency content metric), these same two participants did not have statistically significant differences in any metrics except duration between events. Therefore, Participants 10 and 25 would be good candidates for further investigation to investigate any similarities of their AEs with those of the control group AEs.

In addition, another participant (Participant 27) had statistically significant differences to the overall noise group for all metrics. Furthermore, Participant 53 had statistically significant differences to the overall noise group for all metrics except peak amplitude. However, in contrast to Participants 10 and 25, when comparing the KS tests using the control group data Participant 27 had statistically significant differences in all metrics and Participant 53 had statistically significant differences in all metrics except duration between events. These results suggest that the AEs of Participants 27 and 53 were significantly different to both those of the overall noise revision group (which they were a part of) and the control group. A further KS test between the paring of Participant 27 and Participant 53 also showed statistically significant differences in all metrics. This result indicates that the AEs of these two participants were also significantly different to each other. Therefore, given that the THR implants of both Participants 27 and 53 were in fact revised for noise, the results of the statistical tests confirm the significant variability present in the overall AEs from participants categorised into the noise revision group.

5.4.6 Kolmogorov-Smirnov Tests – Individual Loosening Revisions

KS tests were also performed for the seven metrics using the pairings of the overall empirical CDF of the loosening revision group with the empirical CDF of each loosening revision participant individually. Furthermore, KS tests were performed between the individual loosening revision participants and the overall control group to investigate similarities of the

loosening revision participants to the control group. There were 21 participants who had been categorised into the loosening revision group. Hence, two sets of 21 KS tests were carried out to compare each individual loosening revision group participant to the overall loosening revision and control empirical CDFs. The results of these KS tests are shown in Table B.5 to Table B.8 in Appendix B and a significance level of 1% was used for all the tests.

For the duration metric, 14 of the 21 loosening revision participants had datasets with statistically significant differences to both the overall loosening revision group and overall control group empirical CDFs. This result indicates that two thirds of the loosening group participants had AE events with general duration characteristics that were significantly different to those of the control group. However, the same two thirds of participants also had AE event duration characteristics that were significantly different to those of the overall loosening group. Hence, it was indicated that a significant variation existed in the general event duration characteristics of loosening revision participants. One additional loosening participant had a dataset with a statistically significant difference to only the overall control group dataset. The KS tests for the remainder of the loosening participants failed to show any statistically significant differences in event durations to the loosening group or control group. Therefore, while the KS test results did not make the conclusion that these remaining loosening participants had similar event duration characteristics, similarity in the event duration behaviour of the remaining loosening participants remained a possibility.

For the RMS amplitude metric, five loosening participants had datasets with statistically significant differences to the overall control group empirical CDF. Four of these five participants were observed to be the four loosening participants with statistically significant differences to the overall loosening group empirical CDF for the RMS amplitude metric. For the peak amplitude metric, four loosening participants had datasets with statistically significant

differences to the overall control group empirical CDF. Three of these four participants were observed to be three of the four loosening participants with statistically significant differences to the overall loosening group empirical CDF for the peak amplitude metric. Once again since failing to find a statistically significant difference does not confirm similarity in the distributions, it cannot be concluded that for event amplitudes the majority of the participants had similar behaviour. However, since the results from Section 5.4.3 could not find a statistically significant difference in the event amplitude behaviours of the overall loosening and control group empirical CDFs, the possibility still remains that the majority of the loosening participants have event amplitude behaviours that are similar to both the overall loosening group as well as the control group.

For the duration between events metric the KS test set concerning the overall loosening group had 12 results with statistically significant differences while that concerning the overall control group had 8 results with statistically significant differences. For the rise time metric the KS test set concerning the overall loosening group had 9 results with statistically significant differences while that concerning the overall control group had 11 results with statistically significant differences. In addition, the number of events per recording duration KS test set using the overall loosening group had 7 results with statistically significant differences while that using the overall control group had 14 results with statistically significant differences. Furthermore, similar to the previous KS tests, all but two of the 42 loosening tests had a statistically significant difference for the main frequency content event metric. Thus, further demonstrating the significant variability in the main frequency content of the AEs.

5.4.7 Conclusions from Empirical CDFs and KS Tests

In general, a range of different observations were clear across the different metrics. From the empirical CDF plots and probability density histograms of Section 5.4.2 the overall noise

revision category was the only group that had obvious differences for most metrics which was likely attributed to the presence of audible squeaking AEs. The usefulness of being able to categorise a patient into the noise category based on data that includes audible squeaking is limited as it is likely that both the patient and their surgeon would already be aware that the implant is making audible noises and hence they would not need an AE assessment to produce this diagnosis. However, patients who eventually get revised for implant squeaking may produce non-audible, non-squeaking AEs at an earlier stage that could provide advanced warning of the need for implant revision. Therefore, it may be possible to categorise a patient into the noise revision reason category without their implant producing audible squeaks or other audible sounds.

A major observation from the KS test analyses was that the main frequency content distributions showed statistically significant differences for almost all tests. This result indicates significant variability in the main frequency content of AEs across participants. Therefore, it is unlikely that the main frequency content metric could be used as a diagnostic indicator without incorporating other information about the AEs.

The control group was observed to have an obvious difference to the THR participant groupings as it showed lower values for the number of events per recording duration metric which was consistent with prior observations of the time domain AE recordings. The individual control participants, in general, did not yield a large number of statistically significant differences in the AE metrics (excluding main frequency content) when comparing to the overall control group. However, as stated previously, this lack of statistically significant differences does not automatically confirm general similarity between the control AEs, it simply does not discount similarity.

Furthermore, the KS testing of the noise revision group found that, in general for each metric, over half of the participants categorised into the noise revision group showed evidence that they had AE behaviour that differed significantly from the overall behaviour of the group itself. This result indicates significant intra-participant variation in the AEs of participants who were identified as being in the noise revision group. Additionally, it was concluded that more analysis would be needed to investigate the possibility that some noise group participants had AE content that was more similar to that of the control group AEs than those of participants with known audible AEs.

Overall, the KS tests of the individual loosening revision participants showed varied results between the metrics in terms of statistically significant differences. Some results seemed to suggest that for some metrics there were significant differences between the event characteristics of a considerable number of the loosening participants and the overall loosening group. The results of the KS tests also did not entirely reject the possibility of loosening revision participants having similar event characteristics to the overall control group. Furthermore, for the main frequency content and event duration metrics over half of the loosening revision participants showed statistically significant differences to both the overall loosening group and the control group indicating a large variability in these metrics.

In conclusion, apart from the presence of audible squeaking, it seemed that no major distinctions between the THR revision reasons was present. Ideally, if obvious distinctions were present in the CDFs for each grouping within a metric then it would be likely that the corresponding metric could then be used as a diagnostic indicator for future AE detection. The empirical CDFs of Figure 5.10 showed that a particular event in any grouping could take on almost any value in the range of the particular metric but other events with a similar value of that metric would also have occurred in the other groupings. The implication of this result was

that none of the metrics could be used in isolation to classify an event into a revision reason category as all groupings were likely to contain events with the same metric value of that event. Therefore, the conclusion from the CDF and KS test analysis was that the single metrics alone were not an effective way to categorise participant AE data. However, by comparing or combining multiple metrics more distinct differences between revision reason groupings may be able to be determined and hence more reliable categorisation or diagnosis could be made.

5.5 Multi-dimensional Metric Comparisons

5.5.1 Two-dimensional Event Metric Analysis

The analysis of Section 5.4 concluded that no AE event metric on its own was an effective way to categorise THR AE data. However, more insight may be gained by comparing multiple AE metrics simultaneously. As explained in Section 5.4, each AE event was represented by five signal metrics (duration, rise time, RMS amplitude, peak amplitude, and main frequency content). Two additional metrics concerning the rate of occurrence of events were also computed (duration between events and number of events per recording duration). Since the two additional metrics described the rate of occurrence of AE events they did not correspond directly to specific AE events (i.e. a specific AE event had a "duration" and "rise time" etc. but did not have a "duration between events" or a "number of events per recording duration"). Therefore, only the event duration, rise time, RMS amplitude, peak amplitude, and main frequency content metrics were compared in a multi-dimensional analysis.

Two-dimensional comparisons were performed using scatter plots to provide visual representations of the combined distributions of the two-metric pairs of the five AE event metrics. That is, each AE event signal was represented on a two-dimensional plot where each axes was one of the five AE event metrics (e.g. an AE event which had a duration of 10 ms and

a RMS amplitude of 0.1 V would appear as a single data point located at the corresponding position on a duration and RMS amplitude scatter plot). Ten permutations of two-metric comparisons were possible from the five AE event metrics. Hence, ten different two-dimensional comparisons were performed for the AE event data. Four scatter plots were produced for each permutation of two-metric comparisons in order to compare the AE events across the four different participant data groupings of control, noise revision, loosening revision, and excessive wear revision. Each set of four scatter plots for the ten two-metric comparisons are presented in Figures 5.15-5.23. Note that, like the empirical CDFs of Figure 5.10, due to the data for each metric spanning multiple orders of magnitude, the plots of Figures 5.15-5.23 have been presented using a log scale in order to more clearly show any distinctions.

The scatter plots of Figures 5.15-5.23 also showed the relative density of AE events through the use of a colour bar. The density indicated the relative number of occurrences of AE events within two-dimensional regions of the scatter plot. Consequently, the relative density of the scatter plot expressed the two-dimensional distribution of the data and was analogous to the one-dimensional distributions shown by the magnitudes of the histogram bars in the probability density histograms of Figures 5.11-5.14. For each separate two-dimensional metric comparison (Figures 5.15-5.23), the relative densities were normalised across the four participant grouping scatter plots so that the densities were comparable. A further detail to note was that, prior to normalisation across groups, the densities of each individual scatter plot were divided by the total duration of the AE recordings that contributed to that particular scatter plot. This prior division operation was performed in order to account for the disparity in the total recording durations for the different groupings to avoid a bias in the density values (due to differing number of participants and individual recording durations).

Figure 5.15 shows the scatter plots of the event duration and RMS amplitude for the control, noise revision, loosening revision, and excessive wear revision groups. It was observed from Figure 5.15 that, in general, a slight positive correlation existed in the data. For durations below approximately 10 ms, the RMS amplitude increased as the duration increased (see annotation in Figure 5.15a). However, across all four groups, for durations greater than approximately 10 ms the RMS amplitude values appeared to be uncorrelated to the durations. Lower RMS amplitudes were expected for the very short duration AE events given that RMS amplitudes were calculated using a summation of the squared voltage samples in the signal (i.e. shorter duration AE signals have less voltage samples in the summation). When comparing the scatter plots across the four groups, the noise revision group (Figure 5.15b) had an obvious increased amount of AE events in the uncorrelated region of duration values above 10 ms. Furthermore, a similar trend in the relative densities was observed across the four groups with the maximum densities for each scatter plot occurring in similar regions of the plots. Additionally, the loosening revision group was observed to have the largest relative density values, as indicated by a greater number of regions coloured brown and red (see annotation in Figure 5.15c). The control group had the lowest maximum relative density of the four groups at approximately 0.6 times the maximum density of the loosening group (as indicated by the light green region in Figure 5.15a which corresponded to a relative density of approximately 0.6). This result aligns with expected behaviour that control participants exhibited fewer AE events, leading to the lower density of points.

Figure 5.16 shows the scatter plots of the event duration and main frequency content for the four participant groupings. It was observed from Figure 5.16 that the event durations and main frequency content was largely uncorrelated. However, it was observed that, across all four groupings, the AE events with main frequency content above approximately 5 kHz had durations less than 20 ms. Furthermore, it was evident that AE events with durations above

20 ms generally had main frequency content between 2.5-3.0 kHz. In addition, from Figure 5.16 it was observed that the loosening and excessive wear revision groups had a cluster of events with main frequency content just below 2 kHz which was not evident in the scatter plots of the other groups. The clusters below 2 kHz had low relative density values indicating other regions of the scatter plots had many more occurrences of events. Nevertheless, these clusters below 2 kHz were a unique feature of the AEs of two of the participant groups (loosening and excessive wear) and further analysis of the AE event signals within these clusters may contribute to the diagnostic capability of THR AEs.

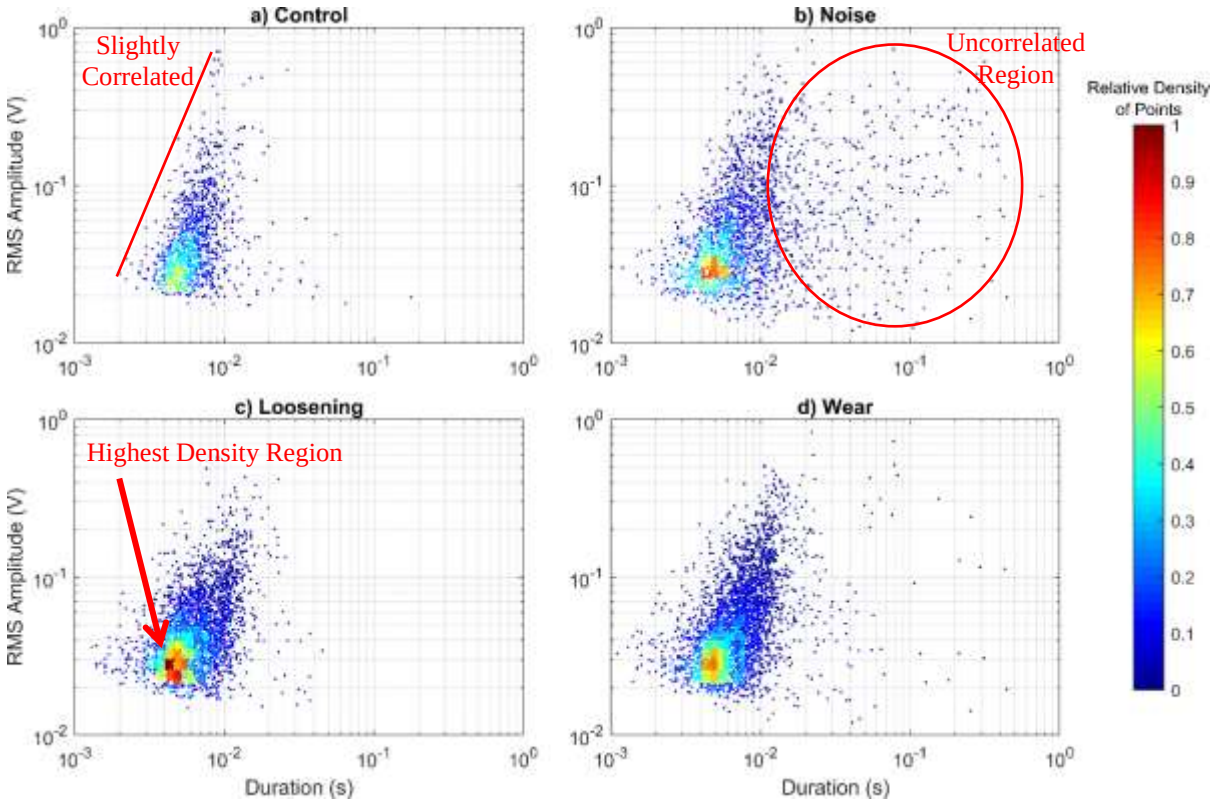

Figure 5.15: Scatter plots of event duration and RMS amplitude for a) the control group, and b) noise, c) loosening, and d) excessive wear revision groups. Note the relative densities were normalised across the four participant grouping scatter plots so that the densities were comparable.

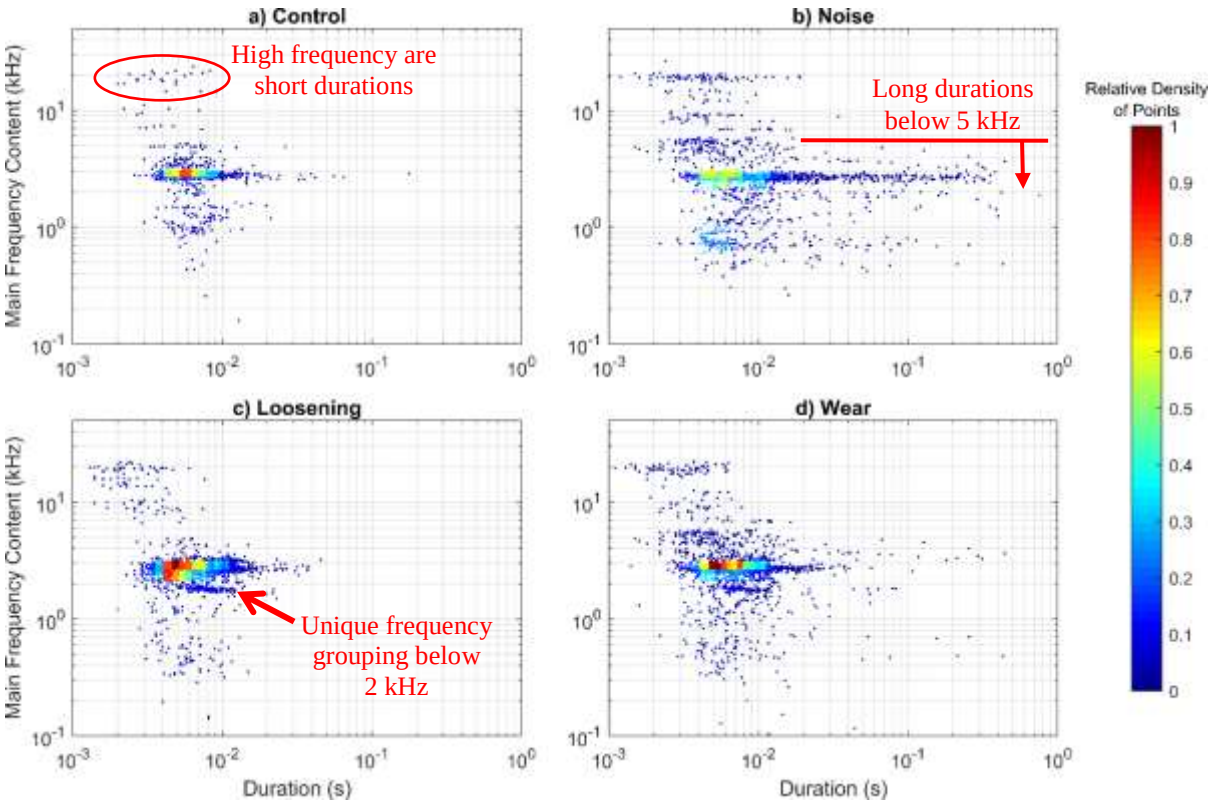

Figure 5.16: Scatter plots of event duration and main frequency content for a) the control group, and b) noise, c) loosening, and d) excessive wear revision groups. Note the relative densities were normalised across the four participant groupings.

Figure 5.17 presents the scatter plots of the RMS amplitude and main frequency content for the four participant groupings. It was observed from Figure 5.17 that the RMS amplitude and main frequency content of the AE events was uncorrelated for all four participant groups. Figure 5.18 shows the scatter plots for the event durations and rise times for the four participant groupings. In contrast to Figure 5.17, it was observed from Figure 5.18 that a moderate positive correlation exists between the event durations and rise times. This correlation is a somewhat expected result given that the maximum rise times could not exceed the duration of an AE event. However, it was possible for long duration AE events to have very short rise times. Despite the possibility of short rise times in long duration AE events, the scatter plots of Figure 5.18 showed that longer duration AE events generally had longer rise times.

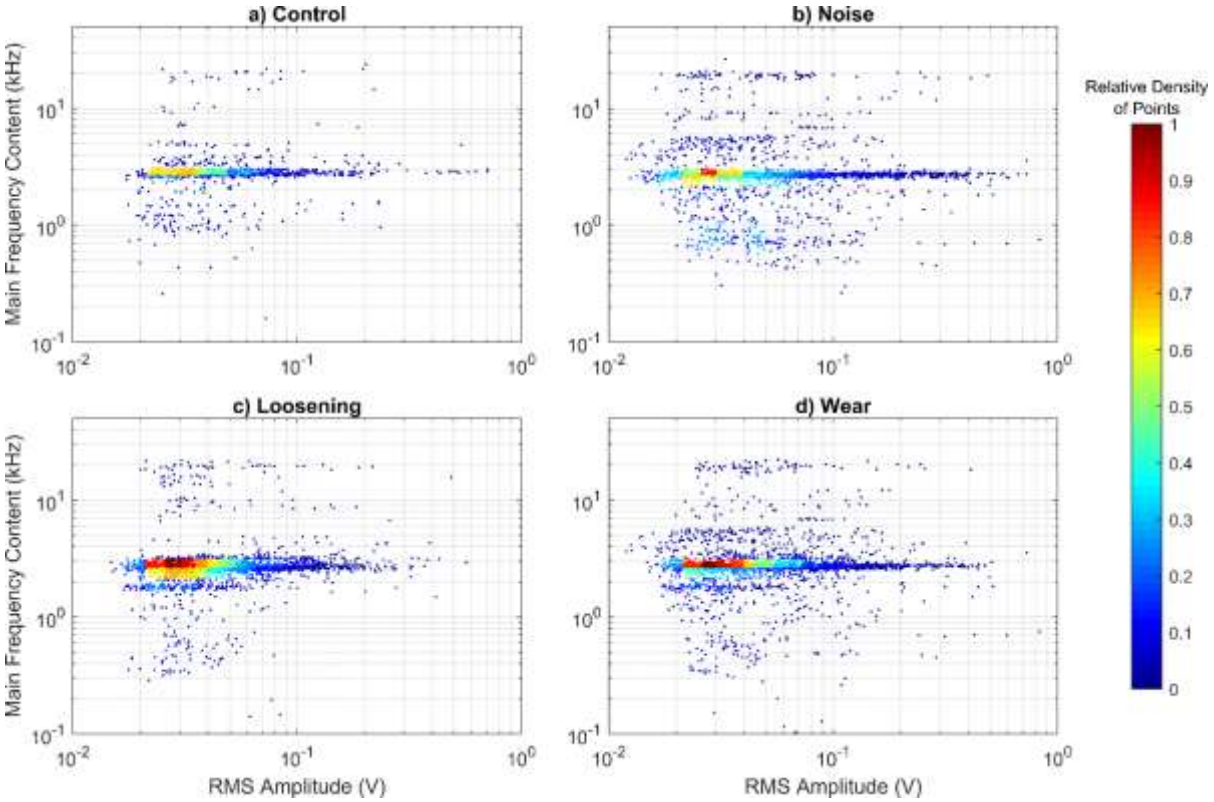

Figure 5.17: Scatter plots of event RMS amplitude and main frequency content for a) the control group, and b) noise, c) loosening, and d) excessive wear revision groups. Note the relative densities were normalised across the four participant groupings.

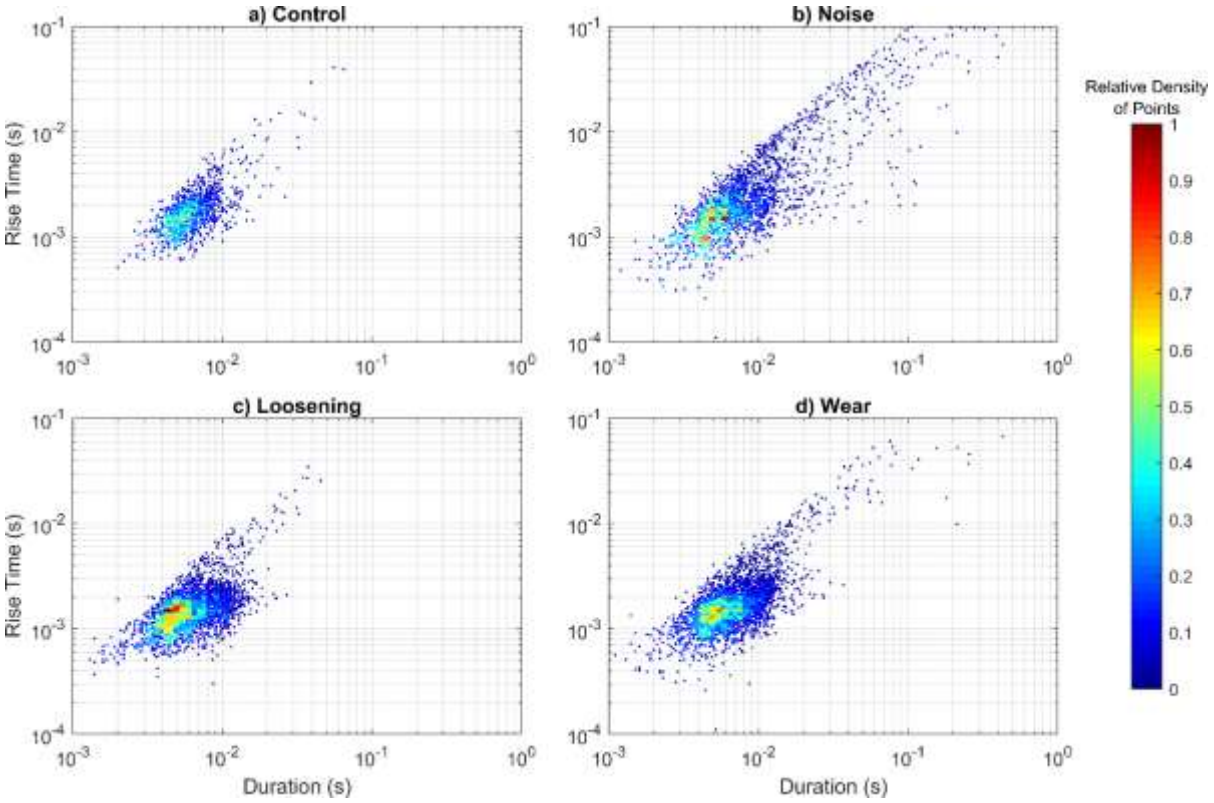

Figure 5.18: Scatter plots of event duration and rise time for a) the control group, and b) noise, c) loosening, and d) excessive wear revision groups. Note the relative densities were normalised across the four participant groupings.

Figure 5.19 shows the scatter plots for the event durations and peak amplitude for the four participant groups. It was observed from Figure 5.19 that the trends for all four groupings were the same as the trends observed for the event duration and RMS amplitude metric combination of Figure 5.15. Figure 5.20 shows the scatter plots for the peak amplitude and main frequency content metric combination for the four participant groupings. It was observed from Figure 5.20 that, much like the metric combination of RMS amplitude and main frequency content in Figure 5.17, the peak amplitudes and main frequency content were uncorrelated for all participant groupings.

Figure 5.21 presents the scatter plots for the event peak amplitudes and RMS amplitudes for the four participant groupings. A relatively strong correlation was observed visually across all four participant groupings for the peak and RMS amplitudes of Figure 5.21. The observed correlation in Figure 5.21 was a generally expected result. It was reasonable to assume that larger peak amplitudes would indicate larger RMS amplitudes given that the peak amplitude contributes to the RMS amplitude calculation and large peak amplitude voltages generally had increased voltage amplitudes surrounding them. However, it was still possible for an AE event signal to have a large peak amplitude voltage with the remainder of the signal having relatively low voltage amplitudes resulting in a low RMS amplitude. Upon closer inspection of the noise revision grouping, a small number of AE events were observed to have relatively low RMS amplitude values and higher peak amplitudes which distinguished them from the main cluster of AE events (see ellipse in Figure 5.21b). These distinguished AE events of Figure 5.21b are likely to be short ‘impulse’ type AE events which can have high peak voltage amplitudes but relatively low RMS amplitudes. However, the general trends of Figure 5.21 indicated that AE events with larger peak amplitudes also had corresponding larger RMS amplitudes.

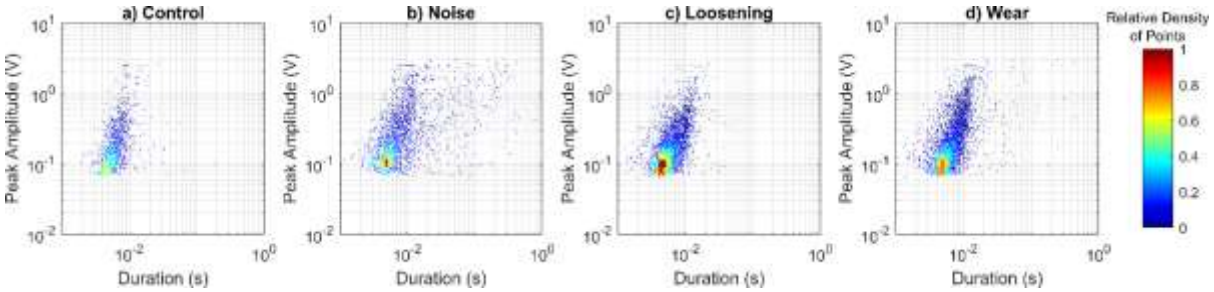

Figure 5.19: Scatter plots of event duration and peak amplitude for a) the control group, and b) noise, c) loosening, and d) excessive wear revision groups. Note the relative densities were normalised across the four participant groupings.

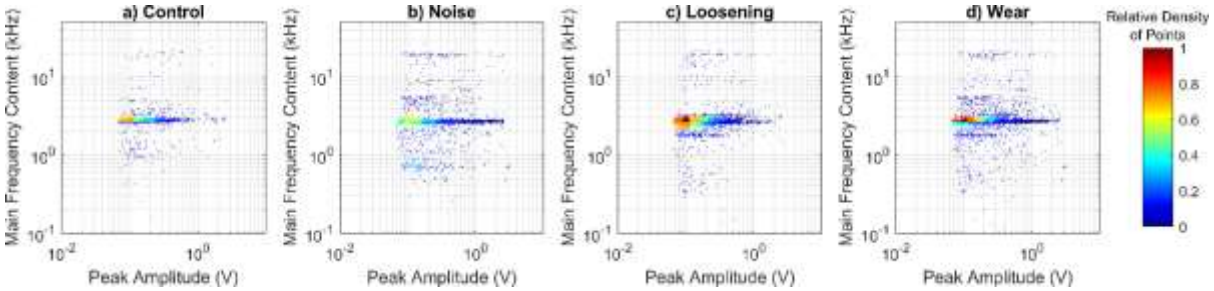

Figure 5.20: Scatter plots of event peak amplitude and main frequency content for a) the control group, and b) noise, c) loosening, and d) excessive wear revision groups. Note the relative densities were normalised across the four participant groupings.

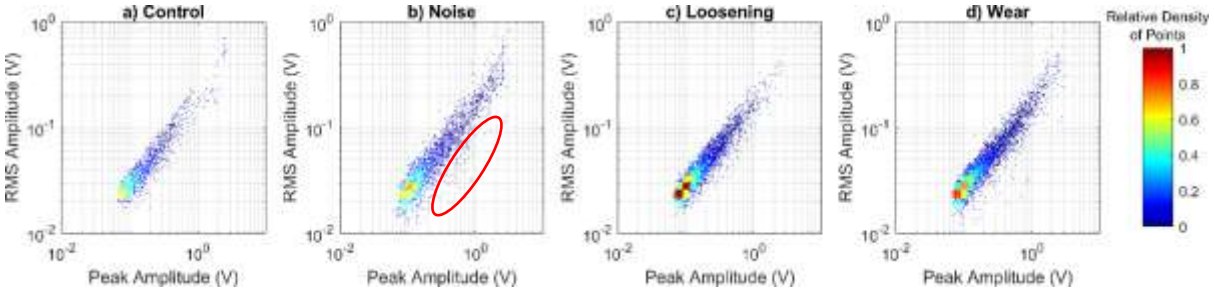

Figure 5.21: Scatter plots of event peak amplitude and RMS amplitude for a) the control group, and b) noise, c) loosening, and d) excessive wear revision groups. Note the relative densities were normalised across the four participant groupings.

Figure 5.22 presents the scatter plots of the event peak amplitudes and rise times for the four participant groupings. It was observed from Figure 5.22 that, for rise times shorter than approximately 1.5 ms, peak amplitude increased as rise time increased. However, for rise times longer than approximately 1.5 ms the metrics were uncorrelated. Figure 5.23 shows the scatter plots of the event RMS amplitudes and rise times for the four participant groupings. From Figure 5.23 trends very similar to those observed in the scatter plots of the peak amplitude and rise time metric combination (Figure 5.22) were observed. The similar trends observed between Figures 5.22 and 5.23 were expected given the correlation that was observed previously between the RMS and peak amplitudes in Figure 5.21.

The final two-metric combination of rise time and main frequency content is presented in Figure 5.24 with one scatter plot for each of the four participant groupings. From Figure 5.24 it was observed that the trends across all the four participant groupings were very similar to those observed in the scatter plots of event duration and main frequency content (Figure 5.16). Comparable to the similarities between Figures 5.22 and 5.23 involving the peak and RMS amplitudes, the similarities observed between the scatter plots of Figure 5.24 and Figure 5.16 were expected given the moderate correlation that was observed between the event durations and rise times of Figure 5.18.

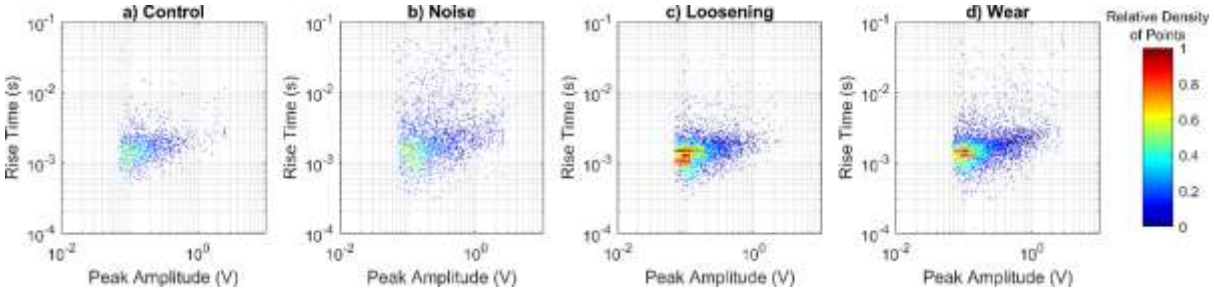

Figure 5.22: Scatter plots of event peak amplitude and rise time for a) the control group, and b) noise, c) loosening, and d) excessive wear revision groups. Note the relative densities were normalised across the four participant groupings.

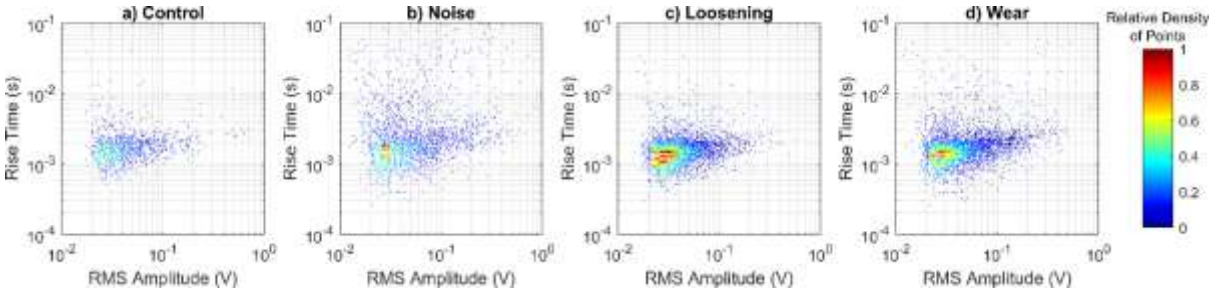

Figure 5.23: Scatter plots of event RMS amplitude and rise time for a) the control group, and b) noise, c) loosening, and d) excessive wear revision groups. Note the relative densities were normalised across the four participant groupings.

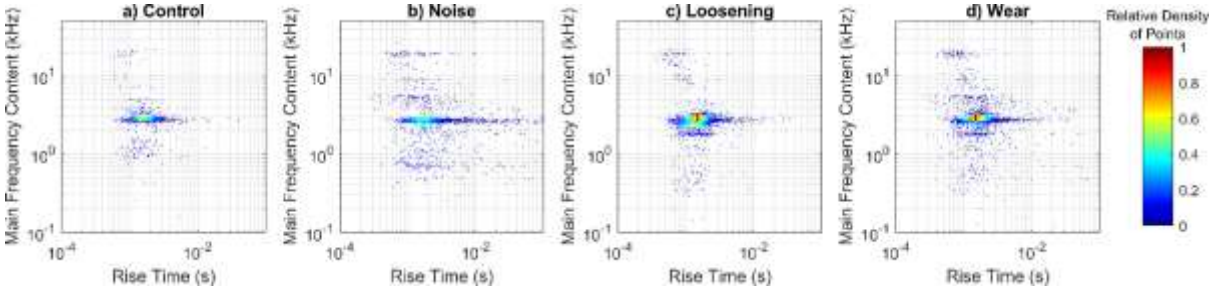

Figure 5.24: Scatter plots of event rise time and main frequency content for a) the control group, and b) noise, c) loosening, and d) excessive wear revision groups. Note the relative densities were normalised across the four participant groupings.

From the two-dimensional scatter plot comparisons of Figures 5.15-5.23 it was found that, in general, one main cluster of AE events was present which had a relatively wide spread much like what was observed in the single metric distributions of Section 5.4.2. Across all the two-metric comparisons, the main data clusters were similar in shape when comparing between the four participant groupings of control, noise revision, loosening revision, and excessive wear revision. However, for all two-metric comparisons the noise revision grouping had additional low relative density AE events that spread away from the main data cluster. With the exception of the additional data spread in the noise revision grouping, the two-dimensional analysis showed that a randomly chosen AE event was likely to have event metrics that were common to all participant categories. Therefore, the main conclusion from the two-dimensional event metric analysis was that it may be possible to categorise a THR AE event into the noise revision group but it was more likely that a random AE event would have event metrics that are common to all participant groupings. However, the two-dimensional analysis was useful in that a number of interesting observations were made that were either not apparent or unable to be noticed from the previous single metric distribution analysis of Section 5.4. Table 5.11 provides a brief summary of the main observations from the scatter plots of each of the two-metric permutations.

Table 5.11: Summary of main observations from 2D scatter plot event metric comparisons.

Metric Pairing	Scatter Plot Visual Observations Summary
Duration – RMS Amplitude	• High RMS amplitudes have longer durations for durations below 10 milliseconds • Uncorrelated for durations greater than approximately 10 milliseconds
Duration – Main Frequency	• Main frequencies above 5 kHz have durations less than 20 milliseconds • Durations above 20 milliseconds mostly have main frequencies between 2.5-3.0 kHz • Loosening and excessive wear revision groups show a unique grouping of events with main frequencies just below 2 kHz
Duration – Peak Amplitude	• Same as Duration – RMS Amplitude
Duration – Rise Time	• An expected and relatively strong correlation observed
RMS Amplitude – Main Frequency	• Uncorrelated
RMS Amplitude – Peak Amplitude	• An expected and relatively strong correlation observed • A very small number of events had low RMS amplitudes but high peak amplitudes which made them appear as outliers to the main data cloud
RMS Amplitude – Rise Time	• Uncorrelated for rise times above approximately 3 milliseconds • High RMS amplitudes have longer rise times for rise times below approximately 3 milliseconds
Peak Amplitude – Main Frequency	• Uncorrelated
Peak Amplitude – Rise Time	• Larger peak amplitudes have longer rise times for rise times below approximately 1.5 milliseconds • Uncorrelated for rise times longer than 1.5 milliseconds
Rise Time – Main Frequency	• Similar to Duration – Main Frequency

5.5.2 Principal Component Analysis

As an extension to the two-dimensional analysis, a five-dimensional analysis that compared all five of the AE event metrics simultaneously was performed on the event data. The implication of performing a five-dimensional analysis was that visual comparisons can only be carried out in up to three dimensions (i.e. 3-D scatter plots). Hence, comparing four or more metrics is

impossible visually without combining the metrics in some way. Therefore, the five-dimensional analysis was performed using a principal component analysis (PCA).

PCA is a multivariate technique that extracts the important information from data that contains multiple dependent variables. The goal of PCA is to express the important information of a dataset as a new set of orthogonal variables, known as principal components (PCs), which are a linear combination of the original variables (Abdi & Williams, 2010). The PCs are computed to express the data in such a way that the first PC is aligned with the direction of largest possible variance in the dataset. Consequently, the first PC is said to 'extract' or 'explain' the largest proportion of the inertia of the dataset (Abdi & Williams, 2010). The inertia of a dataset is analogous to the total variance and is defined as the sum of all the squared elements of each variable (i.e. the sum of the squared distance of each data point from the origin). The second PC is constrained to be orthogonal to the first PC and is orientated such that it extracts the largest proportion of the remaining inertia of the dataset. Subsequent PCs are computed in the same manner until all the inertia of the dataset has been extracted or accounted for. Therefore, the total number of PCs does not exceed the total number of original variables.

Since a PCA has the goal of expressing the important information of a dataset, the PCA will generally be able to express a significant proportion of the total inertia of a dataset with less dimensions than the original variables. That is to say, a PCA performed on a dataset with five dimensions may be able to express in excess of 90% of the total inertia using three PCs instead of five. Consequently, a PCA has the ability to reduce the dimensionality of a dataset and still maintain a good overall representation of the data. Furthermore, the reduction of a five-dimensional dataset down to three dimensions allows a visual representation of the data to be produced using either a two- or three-dimensional plot with the first two or three PCs as the plot axes respectively. In the case of the five-dimensional dataset of THR AE events (using the

same metrics as in Section 5.5.1), PCAs applied to the four participant groupings of AE events did in fact result in dimensionality reductions down to three PCs. These three PCs accounted for approximately 95% of the variation in the datasets. Therefore, visual comparisons were made possible of the five-dimensional AE event data through the use of PCA.

To apply a PCA to a participant AE event dataset, the dataset was first centred on its mean (zero-meaned) and then standardised. The standardisation was performed by dividing the event metric value for each event by the Euclidean norm of the same event metric across all AE events of the dataset (e.g. each duration value of a dataset was divided by the Euclidean norm of all event durations for the entire dataset). The standardisation of the event metrics was an important step in order to ensure that the PCA did not produce biased results. Since a PCA finds PCs that maximise variance in a dataset, variables (AE event metrics) that had generally large numerical values would dominate the PCs (e.g. frequency values varying in the range 1-10 kHz would dominate duration values that vary in the range of 0.005-0.050 seconds). Therefore, standardisation allows an unbiased PCA to be performed without affecting the relative variance of the individual variables.

After the dataset was centred and standardised, the correlation matrix, $\boldsymbol{X}$, was computed. If the dataset contains I AE events and is represented by J variables, then the zero-meaned and normalised dataset is defined as a $I \times J$ matrix $\boldsymbol{A}$, where the rows of $\boldsymbol{A}$ represent each AE event and the columns of $\boldsymbol{A}$ represent the AE event metrics. Therefore, the correlation matrix is defined as:

$$\boldsymbol{X} = \boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A} \qquad \textbf{(5.1)}$$

Where the superscript $\boldsymbol{T}$ indicates the transpose operation. The correlation matrix by definition contains the correlations between the variables of the standardised data. The correlation matrix

has an important useful property in that the eigenvectors of the matrix are orthogonal and they also correspond to the directions of largest variance of the dataset $\boldsymbol{A}$. Hence, the eigenvectors of the correlation matrix are the PCs of the dataset $\boldsymbol{A}$. Additionally, the eigenvalues represent the inertia associated with the eigenvector or PC; the proportion of inertia for a PC is the associated eigenvalue divided by the sum of all the eigenvalues. Thus, to find the eigenvalues and eigenvectors of the dataset, an eigen-decomposition was performed on the correlation matrix using MATLAB. The eigen-decomposition of the correlation matrix is given as:

$$\boldsymbol{X = VSV^T} \qquad \textbf{(5.2)}$$

Where $\boldsymbol{V}$ is a $J \times J$ matrix of which the columns contain the eigenvectors of $\boldsymbol{X}$ and $\boldsymbol{S}$ is a $J \times J$ diagonal matrix containing the eigenvalues of $\boldsymbol{X}$. The eigenvalues in $\boldsymbol{S}$ are ordered from largest in the first column to smallest in the last column. Consequently, the PCs are ordered from first PC in the first column of $\boldsymbol{V}$ to the last PC in the last column of $\boldsymbol{V}$.

A separate PCA was performed on each of the AE event metric datasets for the four participant groupings of control, noise revision, loosening revision, and excessive wear revision AEs. Table 5.12 shows the percentage of inertia explained by each of the five resulting PCs for each of the four datasets. It was observed from Table 5.12 that the percentage of inertia was similar for each PC across the four groups with the average percentages being 46.3%, 30.0%, and 19.4% for the first, second, and third principal components respectively. Furthermore, on average, the first three PCs combined contained over 95% of the total inertia and over 75% of the total was within the first two PCs. Consequently, a large proportion of the variation in the four AE event datasets could be shown using two-dimensional plots of the AE events projected onto the first two PCs.

Table 5.12: Percentage of inertia explained by each principal component.

Principal Component	**1**	**2**	**3**	**4**	**5**
Control (%)	42.2	35.9	19.6	1.3	0.9
Noise (%)	48.9	27.0	19.7	2.6	1.9
Loosening (%)	47.9	27.5	18.4	5.2	1.1
Excessive Wear (%)	46.0	29.4	19.8	3.5	1.3
Average (%)	46.3	30.0	19.4	-	-

In order to transform the individual AE events of each group into the PC coordinate space, factor scores, $\boldsymbol{F}$, were computed. The factor scores are the coordinates of the original dataset projected onto the PCs. Therefore, the matrix $\boldsymbol{F}$ has the same size as $\boldsymbol{A}$ and is defined as:

$$\boldsymbol{F = XV} \qquad \textbf{(5.3)}$$

Where each row of $\boldsymbol{F}$ contains the projected coordinates of one AE event and the columns ordered from first to last represent the PCs ordered from the first to last respectively. It is useful to note here that Equation 5.3 can be used to compute factor scores for additional observations (known as *supplementary observations*) that did not undergo the PCA. That is, the correlation matrix of supplementary observations can be multiplied by the eigenvector matrix, **V**, to yield corresponding factor scores for the supplementary observations. Supplementary observations must undergo the same centering and standardisation operations that were applied to the original dataset. Figure 5.25 shows the PCA factor scores of the AE events projected onto the first two PCs for the four participant groups. It was observed from Figure 5.25 that the general distribution of the data for all groups was broadly similar since all the groups showed one main cluster of data and two noticeable tails. However, the noise and loosening revision groups both exhibited some obvious differences. The noise revision group showed an increased occurrence of highly scattered AE events between the two data tails and in the direction of the first PC (see annotation in Figure 5.25b). The loosening revision group showed a more concentrated group

of AE events situated between -0.1 and 0 on PC2 and close to the PC1 axis (see annotation in Figure 5.25c). Furthermore, given that the two tails that were observed in the data of Figure 5.25 extend away from the origin approximately symmetrically about PC1, it is likely that at least two different variables contribute to the largest variation in the data. However, the factor score representations are limited in their ability to express what physical quantity/original input event metric each PC actually represents, and so the statement that at least two variables contribute to the largest variation of the data cannot be verified from Figure 5.25 alone.

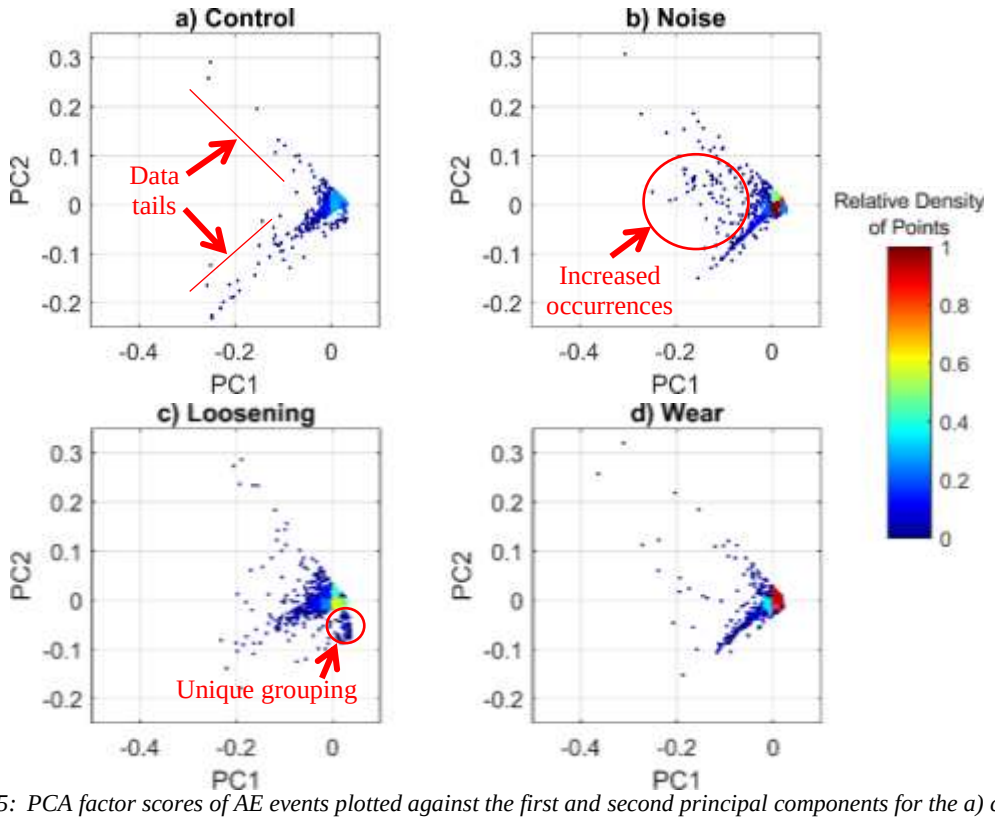

Figure 5.25: PCA factor scores of AE events plotted against the first and second principal components for the a) control, b) noise revision, c) loosening revision, and d) excessive noise revision groups.

An arguably more informative and useful result that comes from a PCA is the correlation between the individual PCs and event metric variables. These correlations provide an estimate for how much information each PC and each event metric variable share. Consequently, the correlations between the PCs and the variables provide a better indication as to the actual meaning (or linear combination) of each PC. Since the eigen-decomposition was performed on

the standardised correlation matrix, the correlation coefficients between the PCs and the variables is defined as:

$$\boldsymbol{L = V\sqrt{S}} \qquad \textbf{(5.4)}$$

Where each row of $\boldsymbol{L}$ represents one of the variables and each column (from first to last) represents a PC from first to last respectively. The correlation coefficients take on values in the range 0-1 and the sum of the squared correlations of each variable add to one (i.e. the rows of $\boldsymbol{L^2}$ sum to one). Therefore, the squared coefficients of the correlations indicate what proportion of each variable is represented by a particular PC. Table 5.13 presents the squared correlation coefficients between each variable and each of the first three PCs for the four participant groupings. Note that the values in Table 5.13 have been multiplied by 100 to display as percentages. Furthermore, the summation of the columns of each of the four separate groups in Table 5.13 will be close to 100% (but not exactly since the values for the fourth and fifth PCs are not shown).

Table 5.13: Squared coefficients of correlation between variables and principal components shown as percentages.

Group	Principal Component	Duration (%)	Peak Amplitude (%)	RMS Amplitude (%)	Rise Time (%)	Main Frequency Content (%)
Control	**1**	56.5	55.7	52.0	46.8	0.0
	2	40.5	41.2	44.3	49.5	4.2
	3	0.3	0.3	0.9	0.9	95.8
Noise	**1**	59.7	64.3	60.0	59.0	1.4
	2	32.0	30.9	35.2	33.1	2.7
	3	1.0	0.0	0.2	1.4	95.9
Loosening	**1**	62.2	79.4	70.1	26.5	0.4
	2	22.6	16.6	22.3	47.0	28.9
	3	0.9	1.1	2.8	17.2	70.0
Excessive Wear	**1**	54.4	70.2	55.4	49.8	0.3
	2	36.4	26.2	40.7	40.4	3.4
	3	0.7	0.1	0.6	1.1	96.3

The major important observation from Table 5.13 was that the duration, peak amplitude, RMS amplitude, and rise time variables dominated the contributions to the first and second PCs across all groups. Conversely, it was evident that the third PC was dominated by the main frequency content for all groups. Therefore, in general, the squared correlation coefficients indicated that the first and second PCs were both primarily linear combinations of the duration, peak amplitude, RMS amplitude, and rise time metrics. Whereas, the third PC primarily represented the main frequency content metric. However, there were some noticeable differences in the percentage values when comparing variables between the four groups despite the general trends being similar.

A visual representation was produced to further interpret the correlations between the variables and the PCs. The correlation coefficients were used as coordinates to plot each variable as a point in the PC coordinate space. Figure 5.26 presents a plot for each of the four participant groupings where the variables have been plotted using the correlation coefficients associated with the first and second PCs (PC1 and PC2). Figure 5.27 presents the equivalent plots for the correlation coefficients associated with the first and third PCs (PC1 and PC3). To aid the interpretation of the variable correlation plots, a unit circle (known as a circle of correlations) was also plotted in the PC coordinate space. As stated previously, the sum of the squared correlation coefficients (coordinates) for a single variable across all PCs is equal to one. Furthermore, recall that the equation defining a unit circle is the sum of the squared coordinates. Consequently, for a two-dimensional representation of the variables in the PC coordinate space, each variable must fall either on or inside the circle of correlations. If the two PCs of the correlation plot can almost completely represent a variable (i.e. if the sum of the variable's squared correlation coefficient percentages for the two PCs is close to 100%) it will be located very close to the circle of correlations.

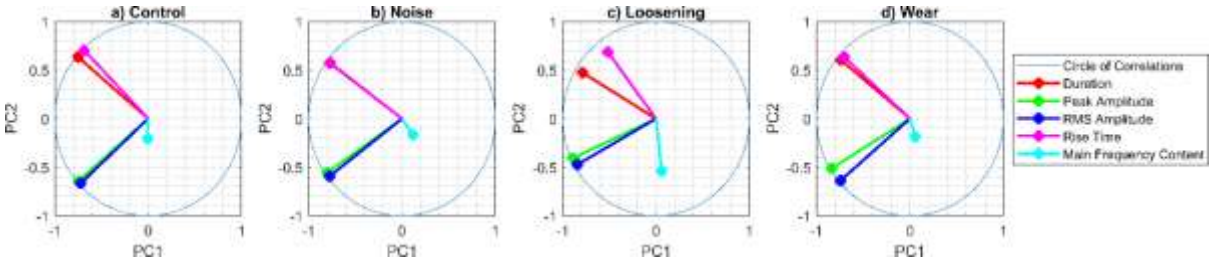

Figure 5.26: Circle of correlations and correlation of variables with the first and second principal components for the a) control, b) noise revision, c) loosening revision, and d) excessive noise revision groups.

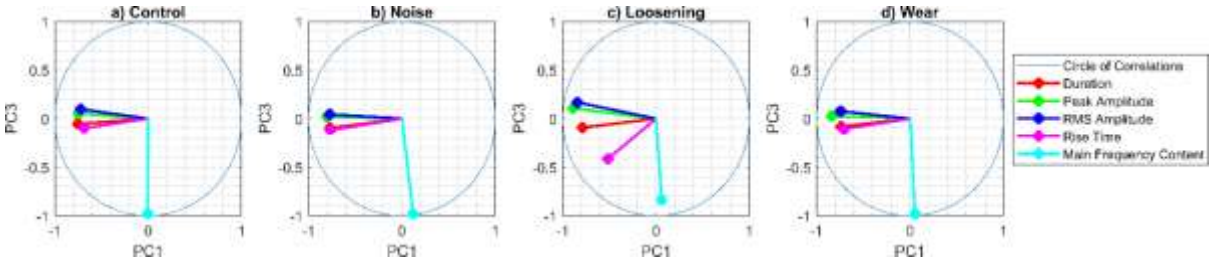

Figure 5.27: Circle of correlations and correlation of variables with the first and third principal components for the a) control, b) noise revision, c) loosening revision, and d) excessive noise revision groups.

It was evident from the squared correlation coefficient results in Table 5.13 that PC1 and PC2 were primarily a linear combination of the duration, rise time, peak amplitude, and RMS amplitude variables. Figure 5.26 confirms the aforementioned observation since the duration, rise time, peak amplitude, and RMS amplitude variables were located close to the circle of correlations. However, the plots of Figure 5.26 offer additional information that was not evident in Table 5.13. The variable correlation plots provide a distinction between positive and negative correlations which the squared correlation coefficients do not. Consequently, a better interpretation of the PCs can be made by considering the plots of Figure 5.26 and Figure 5.27.

It was observed from Figure 5.26 that the duration, rise time, peak amplitude, and RMS amplitude all exhibited very similar negative values with respect to PC1. In this case, the previous interpretation of PC1 remains such that the largest variation in the datasets is a result of similar contributions from the duration, rise time, peak amplitude, and RMS amplitude variables. However, it was also observed from Figure 5.26 that PC2 contrasted the duration and rise time variables with the peak and RMS amplitude variables. The duration and rise time variables took on large positive values while the peak and RMS variables took on large negative

values for PC2. The contrasting of the duration and rise time variables with the peak and RMS amplitude variables explains the presence of the two tails of data observed in the PCA factor scores plots of Figure 5.25. It was observed from Figure 5.27 that the PC3 was almost exclusively defined by the main frequency content variable.

However, a further observation from both Figure 5.26 and Figure 5.27 was that the loosening group had noticeably large differences in the location of the rise time and the main frequency content variables compared to the other groups. The rise time variable for the loosening group was observed to have a greater contribution to PC2 and PC3. Consequently, the rise time variable contributed less to PC1. The main frequency content variable for the loosening group was observed to have a contribution to PC2 that was very similar to the contributions of the peak and RMS amplitudes. Therefore, for the loosening group PC2 was actually a linear combination of all five variables and contrasted the peak amplitude, RMS amplitude, and main frequency content with the duration and rise time variables. Furthermore, PC3 for the loosening group was a linear combination of the main frequency content and the rise time variables.

The variable correlation plots also indicate the correlation between variables. The closer two variables are located to each other on the correlation plot the more correlated the two variables are to each other. It was observed from Figure 5.26 that the peak and RMS amplitude variables were highly correlated to each other with respect to PC1 and PC2. Furthermore, Figure 5.26 shows a high correlation between the duration and rise time variables, with respect to PC1 and PC2, for the control, noise revision, and the excessive wear revision groups. These observations of high correlation between the variables is consistent with the results from the two-dimensional analysis presented in Section 5.5.1. However, it was observed from Figure 5.26 that the duration and rise time variables were less correlated to each other for the loosening group relative to the other groups.

In conclusion, the PCA for the control, noise revision, loosening revision, and excessive wear revision groups showed that the main variation in the datasets was from a combination of the duration, rise time, peak amplitude, and RMS amplitude. Generally, the contribution from the four variables to the main variation (i.e. the first PC) was similar and accounted for an average of 46.3% of the total variation of the datasets. The second PC accounted for an average of 30% of the total variation of the datasets and primarily contrasted the duration and rise time variables with the peak and RMS amplitude variables. The third PC was generally represented by the main frequency content and accounted for an average of 19.4% of the total variation of the datasets. In addition, the results of the PCA demonstrated that the loosening group had some noticeably different characteristics to the other groups.

The PCA has indicated that, while many AEs may be common across the different groups, a definite difference in the distributions of different groups of AEs has occurred, particularly in the loosening group of THR participants. Therefore, the PCA suggests it is likely that particular AEs may be more (or less) prevalent in THR patients with particular implant complications such as loosening. The PCA results add further support for the potential diagnostic capability of the AE monitoring technique. Therefore, further research is needed to identify and examine AEs which may be characteristic to particular implant failure modes.

5.6 Summary

This chapter presented a quantitative analysis of AE events that were detected from the AE signals. The event detection algorithm was shown to provide a good balance between identifying AE events had occurred and adequate distinctions between events that occurred close in time to each other. The subsequent analysis of the metrics computed from the detected events provided an effective way to compare the large amount of detected events across different revision reason participant groups. One of the main conclusions from the AE analysis

of this chapter was that a randomly selected AE event from any study participant would likely have signal metrics that were common to any participant grouping. It was evident in the results of this chapter that the general distributions of the AE event metrics of the control participants were similar to those of the different revision reason groups. Consequently, it was concluded that there was a high likelihood that many of the *in-vivo* THR participant AE recordings contained AE events that might have originated from sources such as other joints in the body, muscle movements, movement or sliding of tendons, fascia, ligaments, and other inherent movement of biological material. The analyses presented within this chapter, in isolation, do not appear sufficient to enable the use of AE sensing as a diagnostic method.

However, the results of this chapter did indicate that some definite differences in the AE event metric distributions of different groups were also present. It was very likely that the *in-vivo* THR AE recordings also contained AE events that occurred only as a result of the presence of the implant. Therefore, it is likely that more insight into failing implants would be gained from identifying the specific types of AE events that caused the differences in the event metric distributions of the THR participant groups. The use of AE as the only sensing method means that only broad conclusions can be made to broadly link AEs to the type of motion the study participant was undertaking at the time, such as walking. The addition of new sensing methods, such as gait tracking of limb angles and implant loads, would provide significant additional insight into the implant mechanics at the time when specific AE signal characteristics are observed. Such additional sensing would enable temporal analyses of recorded AEs, linking signal characteristics to known parameters such as hip articulation angles and approximations of dynamic implant loads. This additional information has the potential to provide significant additional insight and developing a modified sensing system with this potential being the focus of the next chapter.

Chapter 6 Combined Acoustic & Gait Analysis

6.1 Introduction

In this chapter the AE monitoring technique has been combined with gait analysis of the walking motion to provide significantly more insight into the underlying biomechanics of a study participant undergoing AE monitoring. The specific motivation and the basic premise of gait analysis is introduced, followed by an explanation of the basic methodology of the combined test protocol and associated data processing that has been developed. The results from the initial testing of the combined technique performed on three THR participants is presented and discussed.

6.2 Motivation

A limitation of the AE monitoring data presented in Chapter 5 was that no information was available to indicate what stage of the implant motion an AE event occurred. For example, if a squeak occurred during a walking test, there was no data recorded to indicate exactly what stage of the gait cycle the squeak occurred. Additionally, some comparisons of AE event metrics across different participant groups had to be normalised by the total duration of the AE recordings for the group. This normalisation was required to minimise the effect of potential biases introduced by differences in the AE recording times of participants. As stated in Section 3.4.2, the participants performed all test motions in their own time and at self-selected speeds.

The implication of this limitation was that it was impossible to properly investigate the likelihood that some AEs were occurring as a consequence of specific participant (and implant)

motions. A solution to these limitations was to introduce a combined approach of AE monitoring and concurrent gait analysis. Gait analysis is generally considered the quantitative description of all mechanical aspects of walking (Cappozzo, 1984). Through the use of gait analysis, AE events could be associated with specific motions of the hip joint by simultaneously collecting limb motion data of the participant during the AE monitoring. Furthermore, the AE data could ultimately be compared with information such as hip angles and joint loadings through the use of computer modelling. Therefore, gait analysis may provide additional insight into the underlying implant mechanics and help identify cohort-specific AEs. Additionally, the gait analysis will be able to be used select a subset of the AE events that occur during specific stages of the gait cycle. The ability to be able to analyse AE signals in relation to stages of the gait cycle may improve the diagnostic capability of the AE technique.

6.3 Gait Analysis

The typical data that is collected during a gait analysis includes relative orientations and positions of body segments, ground reaction forces, and leg muscle activity (Abernethy et al., 2005; Kadaba et al., 1990). The orientations and positions of the body segments are usually determined using video cameras that track reflective markers attached to the body segments. Ground reaction forces can be measured by walking over a force plate that captures force data in the vertical direction and two orthogonal horizontal directions. Leg muscle activity is usually measured using non-invasive electromyography (EMG) electrodes on the surface of the skin. The initial primary objective of the combined AE monitoring and gait analysis of the THR participants is to associate AEs with the stages of the walking motion and ultimately hip angle data. Therefore, EMG data was not necessary in the gait analysis performed on the THR participants.

As mentioned previously, video cameras are normally used to collect the relative orientations and positions of the body segments during gait analysis which is often referred to as three-dimensional (3D) motion capture (MoCap). 3D MoCap uses multiple video cameras to track the location of reflective markers within a specific volume of space. MoCap can be used to reconstruct the 3D motion of a person if the reflective markers are placed in known locations on the person. The reconstructed motion can be applied to a computer musculoskeletal model to estimate the motion of the limbs and joints using inverse kinematics. Additionally, the computer model can be used to estimate joint loads and muscle forces if a force plate is used in conjunction with the MoCap system to capture ground reaction forces.

Concurrent ongoing research is also being conducted by collaborators to enable the computer modelling to be performed from motion data recorded by multiple inertial measurement units (IMUs) in place of the MoCap system. The ability to perform a gait analysis from IMU data alone will result in a more streamlined data collection process. MoCap systems are relatively expensive and require using specialised equipment in a dedicated space with a limited field of view. Conversely, IMUs can be used as non-invasive wearable sensors that allow for extensive data collection in almost any situation. Nevertheless, the current study presented in this thesis utilises a MoCap system with an incorporated force plate, in conjunction with computer musculoskeletal modelling, to conduct gait analyses on THR participants. The results of the gait analyses presented will be limited to normal walking motions that were performed by the THR participants. The initial focus on walking is an important stepping stone before other more complex motions are considered.

6.4 The Gait Cycle

The gait cycle of normal human walking is generally defined as the time between subsequent heel strikes of one leg. There are a number of key events and phases that characterise the gait cycle. Figure 6.1 presents a schematic of the gait cycle and identifies the key events and phases of the gait cycle. As mentioned, the gait cycle begins with the heel strike of one leg. At approximately 12% of a normal gait cycle the opposite toe is lifted off the ground. The halfway point of a gait cycle (50%) is indicated by the heel strike of the opposite leg. At approximately 62% of the gait cycle the toe is lifted off the ground. And finally, the gait cycle ends (100%) when the subsequent heel strike of the same foot occurs.

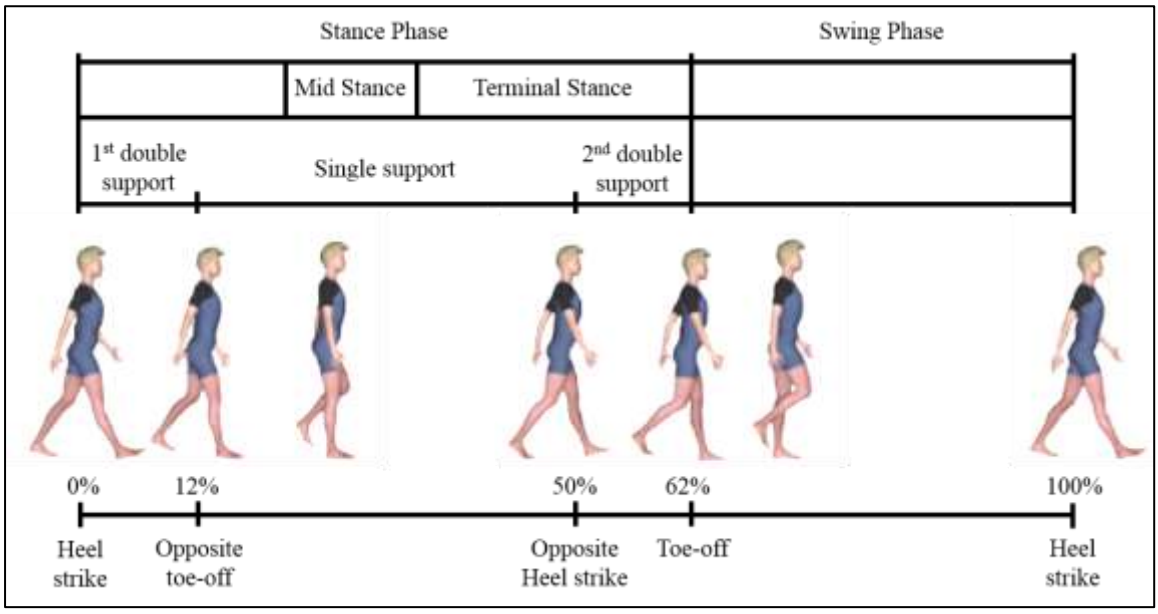

Figure 6.1: The human gait cycle showing key stages of gait. This is a derivative of "The gait cycle" by Iosa et al. (2013) used under CC BY 3.0 licensed by A. J. FitzPatrick under CC BY 3.0.

The gait cycle is also characterised by two main phases: stance phase, and swing phase. The time during which a leg is in contact with the ground is defined as the stance phase of that leg. Therefore, the stance phase commences at 0% of the gait cycle and lasts until toe-off at 62% of the gait cycle. Consequently, the swing phase commences at toe-off and lasts until the heel of that leg strikes the ground again at 100% of the gait cycle. Furthermore, the stance phase

contains three separate stages: the first double support, single support, and second double support stage. The first and second double supports indicate the times when both feet are in contact with the ground (i.e. 0%-12% and 50%-62% of the gait cycle respectively). The single support stage occurs when only one foot is in contact with the ground (i.e. 12%-50% of the gait cycle).

6.5 Methodology

6.5.1 Participants and Equipment

Participants were recruited for the combined AE and gait analysis by the same recruitment process explained in Chapter 3 that was used for the general AE monitoring. A variation to the existing ethical approval (approval number URA/10/11/075, New Zealand Upper South A regional ethics committee) was granted to allow for combined AE and gait monitoring to be carried out. Therefore, prior to commencement of the combined AE monitoring and gait analysis of a participant, an informed consent process was carried out to ensure each participant understood the purpose of the study and what the AE monitoring and gait tracking procedure involved.

The combined AE and gait analysis testing was performed in the biomechanics laboratory at the University of Canterbury where a MoCap system was available. The MoCap system used was a BTS Bioengineering SMART-D Motion Analysis System with six infrared tracking cameras sampling at a rate of 100Hz. A single Kistler 9286BA force plate was used in conjunction with the MoCap system to simultaneously record the ground reaction force of the participant during testing. The sample rate of data from the force plate was also 100 Hz. The AE data collection for the combined testing utilised the same AE equipment and AE testing protocols outlined in Sections 3.2 and 3.4.

In addition to the AE monitoring equipment, three SparkFun 9DoF Razor IMU M0 inertial measurement units (IMUs) were attached to the participant to record motions of the pelvis, femur, and tibia respectively. Each IMU recorded 3D acceleration, angular velocity, and magnetic field strength data at 100 Hz. As mentioned previously, ongoing research is focussed on using the information from the IMUs to perform the gait analysis without the need for the MoCap system. Therefore, in the current study IMU data has been collected but has not directly been used in the estimation of hip joint angles. However, some IMU data was used to synchronise the AE data with the MoCap data, using a fourth IMU sensor synchronised to force plate data through impulse/drop tests at the beginning of the tests.

The specific computer musculoskeletal modelling software that was used for the gait analysis was OpenSim. OpenSim is an open source software program that enables modelling and analysis of the neuromusculoskeletal system (Delp et al., 2007). OpenSim can utilise experimental MoCap and force plate data to analyse the dynamic motions of study participants. The software allows for a generic human skeletal model to be scaled appropriately to match the anthropometric measurements of the test participant. Furthermore, OpenSim employs inverse kinematics to resolve the joint angles associated with the observed motions.

6.5.2 Motion Capture Data Collection

The MoCap data collection procedure involved four key steps: the recording of anthropometric measurements of the participant, the placement of the reflective markers on the participant, the recording of a static trial, and the recording of dynamic trials. This section briefly outlines the four key steps. More detailed and specific instructions for the MoCap data collection can be found in Appendix C and Appendix D. The anthropometrics that needed to be measured and recorded are outlined in Table 6.1. The purpose of the basic anthropometric measurements was

to assist subsequent scaling of the bones of the OpenSim musculoskeletal model which increased the accuracy of three-dimensional motion reconstruction.

Table 6.1: Required anthropometric measurements.

Name	Description
Body Mass	Participant mass
Height	Participant height
Inter-ASIS Distance	Distance between the left anterior superior iliac spine (ASIS) and right ASIS. Note, this only needs to be measured if markers cannot be placed directly on the ASIS (e.g. in obese patients).
Leg Length (Left & Right)	Full leg length, measured between the ASIS marker and the medial malleolus, via the knee joint. Measure with patient standing, if possible.
Knee Width (Left & Right)	The medio-lateral width of the knee across the line of the knee axis. Measure with patient standing, if possible.
Ankle Width (Left & Right)	The medio-lateral distance across the malleoli. Measure with patient standing, if possible.

Reflective markers were placed on the participant in predetermined anatomical positions to allow the MoCap system to track body positions. The marker set used was a modified version of that which is included with the OpenSim *Gait 2392* model. The *Gait 2392* marker set has been developed and refined primarily from previous biomechanical studies which used three-dimensional motion capture (Cappozzo et al., 1995; Davis et al., 1991; Hamner et al., 2010; Kadaba et al., 1990). Collectively, the model utilised 39 markers across the lower limbs, pelvis, torso, and head. For the testing performed in the current study, only 29 of the model markers were used, with two of these 29 markers being optional. Table 6.2 outlines the markers that were placed on anatomical landmarks. It was important to place the markers as accurately as possible on the appropriate anatomical landmarks as disparities in positioning adversely effected the results of the OpenSim modelling process. For the initial testing, a trained physiotherapist assisted with the marker placement to ensure accurate placement.

Table 6.2: Anatomical landmark marker positions.

Definition	**Marker Position** (***** indicates required for static test only)
Left ASIS	Left anterior superior iliac spine.
Right ASIS	Right anterior superior iliac spine.
Sacral	On the skin mid-way between the posterior superior iliac spines (PSIS) and positioned to lie in the plane formed by the ASIS and PSIS points.
Left Knee Lateral	***** On the lateral epicondyle of the left femur and positioned to lie along the estimated axis of rotation of the knee joint.
Left Knee Medial	***** On the medial epicondyle of the left femur and positioned to lie along the estimated axis of rotation of the knee joint.
Left Ankle Lateral	***** On the lateral malleolus of the left shank along the line that would pass through the transmalleolar axis.
Left Ankle Medial	***** On the medial malleolus of the left shank along the line that would pass through the transmalleolar axis.
Left Toe Tip	If wearing shoes, on top of the shoe at the distal end and halfway between the first and second phalanges. If not wearing shoes, over the distal end of the second metatarsal.
Left Heel	On the posterior-most point of the calcaneus (or the shoe) and at the same height above the plantar surface of the foot as the *Left Toe Tip* marker.
Right Knee Lateral	***** Same as *Left Knee Lateral* but on the right femur.
Right Knee Medial	***** Same as *Left Knee Medial* but on the right femur.
Right Ankle Lateral	***** Same as *Left Ankle Lateral* but on the right shank.
Right Ankle Medial	***** Same as *Left Ankle Medial* but on the right shank.
Right Toe Tip	Same as *Left Toe Tip* but on the right foot.
Right Heel	Same as *Left Heel* but on the right foot.
Left Acromium	On the left acromio-clavicular joint.
Right Acromium	On the right acromio-clavicular joint.

Table 6.3 outlines the markers that were placed as part of a three marker cluster on either the thigh or shank of the participant. The purpose of the marker clusters was to allow the upper and lower legs to be tracked in three-dimensions without the need for other markers on the same segments. The exact positioning of the marker clusters was not critical provided that data from a static test was collected first that contained the knee and ankle markers for reference. The model scaling process of OpenSim used the relative positions of the anatomical markers

to accurately position the marker clusters on the model. Therefore, after model scaling, the marker clusters were able to be used to accurately track the position of the corresponding limb. Figure 6.2 shows a study participant ready for testing in the biomechanics laboratory with the markers, AE sensors, and IMUs attached.

Table 6.3: Limb cluster marker positions.

Definition	Marker Position
Left Thigh Upper	The cluster should be placed so that this marker lies on the lateral thigh and approximately in the coronal plane of the femur. Note that the cluster can be positioned at any distance distally along the femur that is practicable.
Left Thigh Front	The cluster should be placed so that this marker lies anterior and distal to the *Left Thigh Upper* marker.
Left Thigh Rear	The cluster should be placed so that this marker lies posterior and distal to the *Left Thigh Upper* marker.
Left Shank Upper	The cluster should be placed so that this marker lies on the lateral shank and approximately in the coronal plane of the tibia. Note that the cluster can be positioned at any distance distally along the shank that is practicable.
Left Shank Front	The cluster should be placed so that this marker lies anterior and distal to the *Left Shank Upper* marker.
Left Shank Rear	The cluster should be placed so that this marker lies posterior and distal to the *Left Shank Upper* marker.
Right Thigh Upper	Same as the *Left Thigh Upper* but on the right thigh.
Right Thigh Front	Same as the *Left Thigh Front* but on the right thigh.
Right Thigh Rear	Same as the *Left Thigh Rear* but on the right thigh.
Right Shank Upper	Same as the *Left Shank Upper* but on the right shank.
Right Shank Front	Same as the *Left Shank Front* but on the right shank.
Right Shank Rear	Same as the *Left Shank rear* but on the right shank.

Figure 6.2: Images of participant with reflective markers, AE sensors, and IMUs attached ready for testing.

A static test was carried out once all markers had been placed on the participant. The static test involved recording marker position data of the participant standing stationary in an upright position on the force plate, near the centre of the calibrated volume. The knee and ankle markers were then removed from the participant once a static test had been performed (See the entries of Table 6.2 marked by a * for markers that were removed following the static trial). Removing the knee and ankle markers simplified the post-processing of the dynamic MoCap data. At least five walking tests of dynamic data were then carried out with each participant.

The walking tests involved recording marker position data at 100 frames per second concurrently with the AE and IMU data as the participant walked across the room. The participant was asked to walk in a straight line and at a self-selected speed. The limited size of the testing laboratory meant that only three consecutive walking gait cycles could be recorded by the motion capture during each test. Furthermore, the participants walked in such a way that the foot of the leg being acoustically monitored contacted the force plate during the second of

the three gait cycles. Therefore, ground reaction force data was captured at 100 Hz for the stance phase of the second recorded gait cycle.

6.5.3 Motion Capture Data Processing and Joint Angle Estimation

For each recorded data frame, the MoCap data from each walking test contained 3D coordinates of each of the 29 markers. The origin of the coordinate system was located at one corner of the top surface of the force plate and the coordinate values had units of metres. During testing, each marker must be in view of at least two cameras in order for its coordinates to be recorded. While it was generally uncommon, sometimes data frames had missing coordinates for particular markers if the marker was obscured at some point during testing. Markers can be sometimes momentarily obscured by things such as loose clothing or the cables from the AE sensors. In any cases of missing marker data, 3D spline interpolation using the marker coordinates from multiple surrounding frames was employed with the BTS MoCap software to estimate values of any missing coordinates. Once any missing markers data had been interpolated, the OpenSim data processing could be performed.

The musculoskeletal model used was the OpenSim *Gait 2392* model. The *Gait 2392* model is a 23 degree of freedom 3D musculoskeletal model of the human torso and lower limbs which also includes a representation of 76 muscles. Figure 6.3a shows an anterior view of the OpenSim *Gait 2392* model with the standard marker set of the model. The first step of the OpenSim processing was to scale the musculoskeletal model to match the specific participant. The scaling process involved using the static trial experimental data to resize each bone and then to shift the model into the pose the participant had during the static trial. To perform the scaling, OpenSim first used the distances between the experimental anatomical markers to rescale the size of each bone individually. After rescaling the bones, the model was fit to the

experimental data in such a way to minimise the distance between each experimental marker and its corresponding marker on the model. Figure 6.3b shows a scaled model fitted to the experimental markers of a static trial. The markers of the model are shown coloured pink in Figure 6.3b and the experimental markers are coloured blue. The cluster markers were not used at all during the scaling process (but still appear as experimental markers in Figure 6.3b) since their absolute positions relative to the musculoskeletal system were unknown at that point of the analysis.

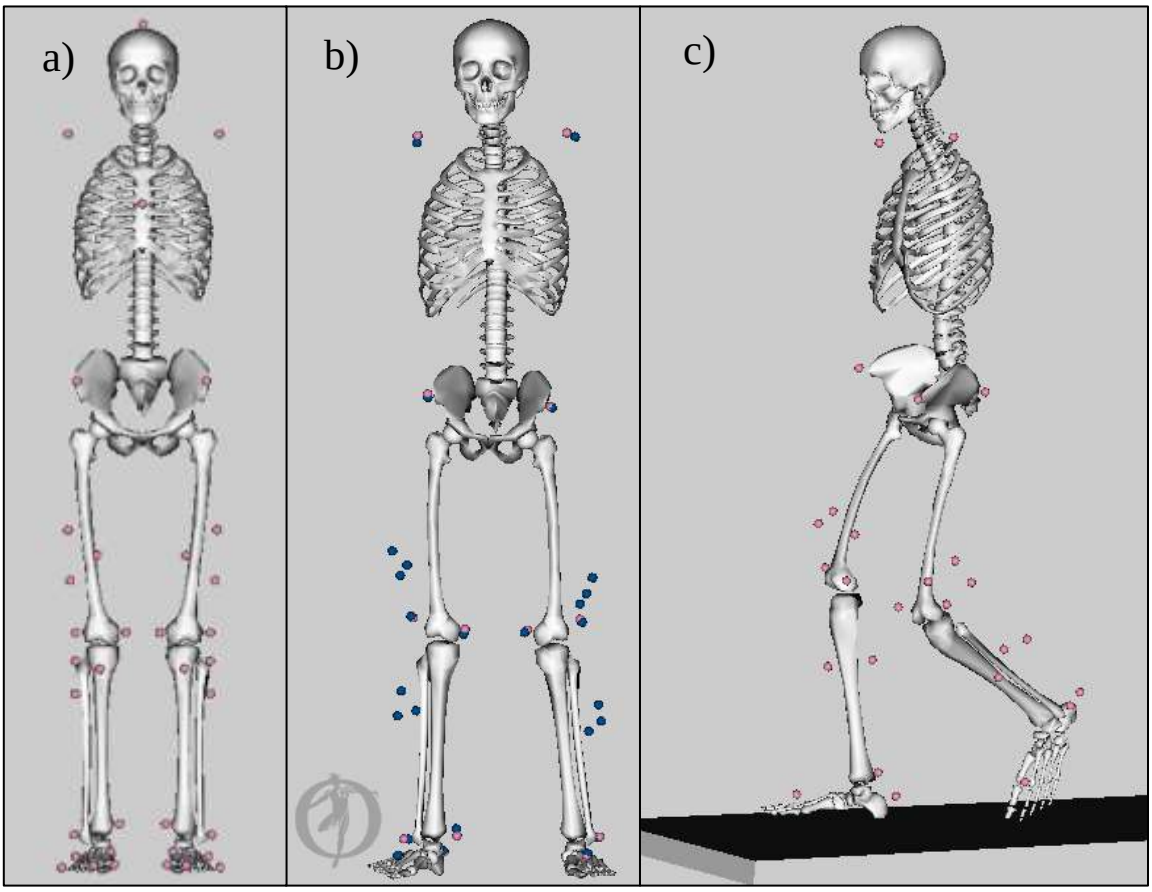

Figure 6.3: OpenSim Gait2392 model showing a) the original model with marker set, b) scaled model with marker set (pink markers) and static trial experimental markers (blue markers), and c) scaled model during walking test.

The OpenSim documentation specified that a satisfactory scaling result was reached when the root-mean-squared (RMS) error in the distances was less than one centimetre and the maximum error was less than two centimetres. Sometimes the positions of particular markers on the model may not represent the position of the experimental marker on the test participant very well.

Consequently, the resultant error from the scaling process would exceed the specified limits. To reduce the error in the scaling process, the position of the markers on the model can be adjusted, if necessary, to more accurately represent the position of the corresponding marker that was attached to the test participant. The scaling was performed again once any model markers had been adjusted and the new error values were checked. This scaling process was iterated until the error values no longer exceed the specified limits.

Once the errors from scaling were below the specified limits, and with the model in the static position, the model markers were then shifted to the locations indicated by the corresponding experimental markers. This step occurs under the assumption that, after scaling, the scaled and shifted model exactly represents the position of the participant during the static trial. Therefore, the positions of all the experimental markers from the static trial should be the locations of their corresponding model markers. Consequently, any model markers that were not in exactly the same position as their corresponding experimental marker are shifted, relative to the model, to the exact location of the experimental marker. Additionally, the locations of the cluster markers relative to the model are known at this point since the corresponding model markers are shifted to the locations of the experimental markers. Note that since it was unlikely that the model was scaled and shifted to exactly match the static trial exactly (i.e. maximum error value was unlikely to be zero centimetres) the locations of the shifted model markers are in fact approximations to within the associated error values.

Finally, to perform the inverse kinematics on the walking tests, OpenSim takes the experimental data from the walking tests and fits the markers of the scaled model to each data frame of coordinates. The OpenSim documentation specified a good inverse kinematics result will have maximum errors less than 2-4 centimetres and an RMS error below two centimetres. Figure 6.3b shows the position of the scaled model for a single frame of a walking test after

the inverse kinematics was applied. Once the inverse kinematics had been performed, the estimated joint angles were output from the resultant model positions. The inverse kinematics result yields the estimated joint angles for each joint of the model. For the hip joint, three sets of angles result from applying the inverse kinematics: flexion, adduction, and rotation. Consequently, since each data frame had an associated timestamp that was also synchronised with the AE and force plate data, the joint angles were able to be plotted against the corresponding AE signal and ground reaction force.

6.6 Results & Discussion

6.6.1 Participant Data

Combined AE and gait data has been collected for three THR participants. Table 6.4 presents the body weight, height, sex, leg acoustically monitored during testing, and implant bearing type for the three participants. All three participants were recruited for this testing because they had indicated that their THR implant often made audible noises. All three participants had an implant with a ceramic-on-ceramic (CoC) main bearing interface.

Table 6.4: Participant sex, body weight, height, leg monitored, and implant type.

Study Number	Sex	Body Weight (N)	Height (cm)	Leg Monitored	Implant Bearing Type
Participant118	F	736	173	Left	Ceramic-on-ceramic
Participant119	M	873	178	Right	Ceramic-on-ceramic
Participant120	M	863	170	Left	Ceramic-on-ceramic

6.6.2 Observations from Participant120 – Implant Squeaking

Of the three THR participants that underwent combined AE monitoring and gait analysis, one participant, Participant120, produced regular audible squeaking during testing. Participant120

performed ten separate normal walking tests that each contained three consecutive gait cycles. Figures 6.4 and 6.5 show the synchronised AE, ground reaction force, and hip angle data for two separate walking tests of Participant120. The horizontal axes of Figures 6.4 and 6.5 show the percentage of gait cycle corresponding to the leg that was acoustically monitored. Since each walking test contained three consecutive gait cycles, the gait cycle percentages of 0%, 100%, and 200% correspond to the start (heel strike) of the first, second, and third walking gait cycles respectively. The second gait cycle (100%-200%) corresponded to the participant stepping on the force plate with the acoustically monitored leg (left leg for Participant 120). A further point to note was that the average velocity of the centre of mass of the participant (estimated through the OpenSim inverse kinematics) was 1.15 ms^{-1} and 1.24 ms^{-1} for the tests of Figures 6.4 and 6.5 respectively.

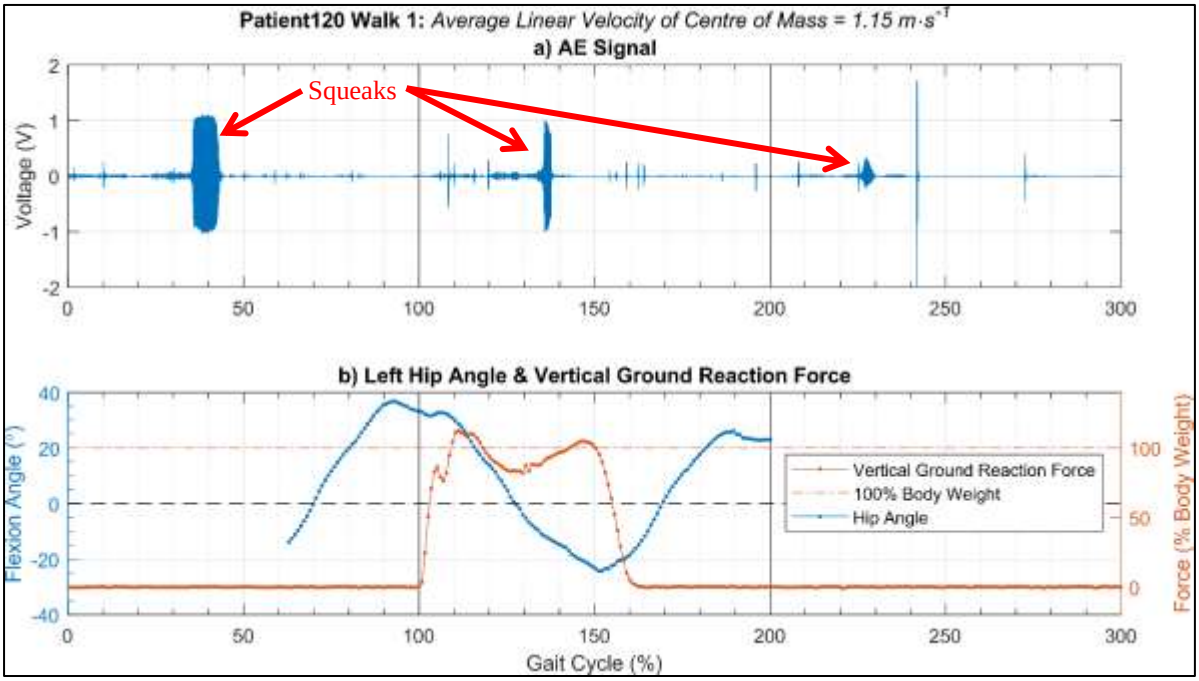

Figure 6.4: a) AE signal from three consecutive gait cycles of a walking test of THR Participant120 with b) synchronised hip angle and ground reaction force data. Note that ground reaction force data was only available for the second consecutive gait cycle.

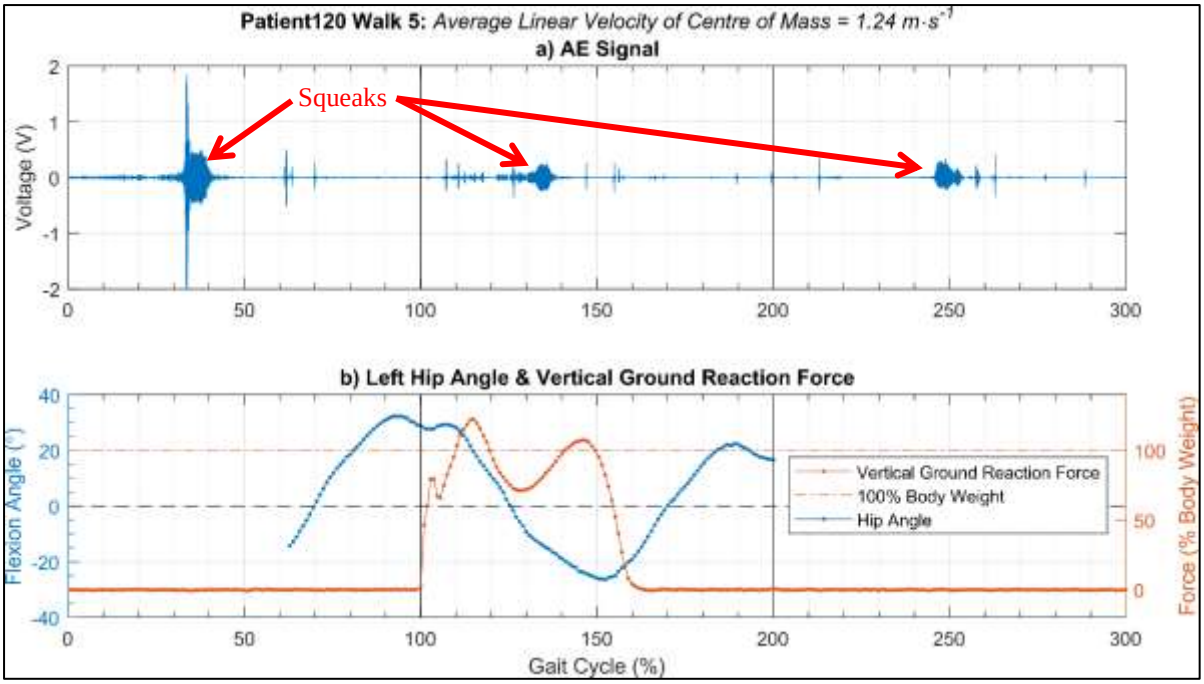

Figure 6.5: a) AE signal from three consecutive gait cycles of a second walking test of THR Participant120 with b) synchronised hip angle and ground reaction force data. Note that ground reaction force data was only available for the second consecutive gait cycle.

It was observed from Figures 6.4a and 6.5a that six squeaks in total occurred between approximately 30%-50%, 130%-140%, and 225%-255% points in the gait cycle across the two tests. The squeaks were indicated by regions of AE signal with sustained increases in signal magnitude (see annotations in Figures 6.4a and 6.5a) as was demonstrated in Section 4.3.2. All six observed squeaks from Figures 6.4 and 6.5 had a different time domain appearance. Nevertheless, it was apparent that all the squeaks occurred during a similar stage of the gait cycle. Figure 6.1 showed that the single support stage of the gait cycle is in the range of 12%-50%. Therefore, the squeaks of Figures 6.4 and 6.5 all occurred during the second half of the single support stage of the gait cycle, also known as terminal stance.

The observation that the squeaks occurred during terminal stance was also observed in all other squeaks that occurred during the walking tests of Participant120. Table 6.5 presents the approximate gait cycle percentage that each observed squeak commenced for all ten walking tests. The percentages of Table 6.5 were approximated to the nearest 5%. A missing entry in Table 6.5 indicates that a squeak did not occur during that particular gait cycle of a test. It can

be seen from Table 6.5 that a squeak occurred during 21 of the 30 gait cycles that were recorded across the ten walking tests. Sixteen of the 21 squeaks occurred during either the first or second gait cycle of their corresponding test. These sixteen squeaks all commenced within the range of 30%-40% of their respective gait cycle with an average of 36%. The remaining five of the 21 squeaks occurred during the third gait cycle of their corresponding test. The squeaks of the third gait cycles commenced within the range of 25%-50% of their respective gait cycle with an average of 41%.

Table 6.5: Approximate percentage of gait cycle that squeak commenced for all walking tests of THR Participant120. The percentages were approximated to the nearest 5% and if no squeak occurred for a particular gait cycle then no percentage is shown.

Test Number	Approximate Percentage of Gait Cycle when Squeak Commenced (%)		
	Squeak 1	Squeak 2	Squeak 3
1	35	135	225
2	30	135	-
3	40	135	-
4	35	140	245
5	35	130	245
6	40	135	240
7	40	-	250
8	30	135	-
9	-	135	-
10	-	-	-
Average	**36**	**136**	**241**
Range	**10**	**10**	**25**

The variability that was observed in the occurrence of squeaks during the third gait cycles of the tests was attributed to the testing methodology. The limited size of the testing laboratory meant that only three full gait cycles could be captured by the MoCap equipment. The participants were required to stop walking almost immediately after they had completed the third walking gait cycle. Consequently, during most walking tests, the participants' third gait

cycle did not resemble a typical walking gait since they had started to decelerate before completing the cycle. Furthermore, in some cases the participants did not complete the final swing phase but instead brought their feet together and stopped walking. As a result, the gait cycle percentages from at least 250% onwards may not be accurately aligned with the AE data. Nevertheless, overall, the squeaks that occurred during the walking tests, especially those of the first and second gait cycles, showed good repeatability with respect to the stage of the gait cycle during which they occurred. This result is an important finding since it implies that the occurrence of squeaking in THR implants is likely related to specific motions of the implant.

The hip angle data of Figures 6.4b and 6.5b was observed to exhibit behaviour in agreement with that of a typical walking gait. The maximum extension of the hip joint during a typical walking gait generally occurs at 50% of the gait cycle (Whittle, 2014), coinciding with the heel strike of the opposite leg. The maximum flexion of the hip joint during a typical walking gait is reached towards the end of the swing phase at approximately 80-90% of the gait cycle (Whittle, 2014). Figures 6.4b and 6.5b show the hip flexion angle to range between approximately -25° (25° extension) at opposite heel strike and approximately 34° (flexion) at 90% of the gait cycle. It was also observed from Figures 6.4b and 6.5b that the hip transitioned from flexion to extension (i.e. positive to negative flexion angles) at approximately 25% of the second gait cycle.

Hip angles during the gait cycle are generally considered to be subject-specific and dependent on multiple factors. However, hip angles during normal walking reported in literature were between maximum flexion angles of approximately 30°-40° and maximum extension angles of approximately 5°-20° (Abernethy et al., 2005; Beaulieu et al., 2010; Bennett et al., 2008; Chung & Wang, 2010; Harris & Wertsch, 1994; Whittington et al., 2008). Therefore, the observed maximum hip flexion angles from Participant 120 of approximately 34° flexion were within

the range of maximum flexion angles reported in literature. However, the observed maximum extension angles from Participant 120 of approximately 25° extension were 5° greater than the maximum values reported in the literature. The large hip extension values that were observed from Participant 120 were possibly the result of the participant subconsciously modifying their gait to ensure they contacted the force plate during testing. Additionally, the hip angle estimates presented in Figures 6.4b and 6.5b were computed through the inverse kinematics analysis in OpenSim. Therefore, it is possible that further refinement of the musculoskeletal model may result maximum hip extension angles that fall within the ranges reported in the literature. Nevertheless, the overall hip angle profiles observed during the entire gait cycle from Participant 120 broadly agreed with overall hip angle profiles found in literature (Abernethy et al., 2005; Beaulieu et al., 2010; Bennett et al., 2008; Harris & Wertsch, 1994; Whittington et al., 2008).

The ground reaction forces of Figures 6.4b and 6.5b were observed to exhibit the same general behaviour to each other. The ground reaction forces began to increase rapidly after second heel strike (0% of the second gait cycle) to a maximum force peak at approximately 15% of the second gait cycle. A momentary minor decrease in the ground reaction forces occurred at approximately 5% of the second gait cycle. After the ground reaction forces had reached their maximum, a decrease occurred to a local minimum at approximately 30% of the second gait cycle. Finally, the ground reaction forces increased to another peak at approximately 45% of the second gait cycle before returning to zero at toe-off (62% of the second gait cycle). Table 6.6 presents a summary of the ground reaction force magnitudes for the walking tests of Figures 6.4 and 6.5. The magnitudes of the ground reaction forces in Table 6.6 have been normalised and are shown as percentage body weight of the participant obtained during the static MoCap trial. The ground reaction forces from the eight other walking tests of Participant120 were

observed to be consistent with those of Figures 6.4b and 6.5b with respect to their overall magnitude trends.

Table 6.6: Summary of ground reaction force for walking tests of Figure 6.4 and Figure 6.5.

Event Name	**Approximate Percentage of Second Gait Cycle**	**Approximate Ground Reaction Force** (% Body Weight)	
		Walking Test 1 (Figure 6.4) Average Speed = 1.15 ms^{-1}	**Walking Test 5 (Figure 6.5)** Average Speed = 1.24 ms^{-1}
Heel strike	0%	0	0
Max/Peak 1	15%	112	122
Min after Peak 1	30%	83	71
Peak 2	45%	105	107
Toe-off	62%	0	0

The behaviour of the ground reaction forces observed from the participant testing was consistent with the behaviour of ground reaction forces of normal walking gait found in literature (Watkins & Mathieson, 2009; Whittle, 2014). The first characteristic peak observed from the ground reaction force was a result of the upward acceleration of the body as the leg took the full load of the body during the beginning of the stance phase. During mid-stance the upward acceleration was reduced as the centre of mass of the body travelled over the leg and the heel began to lift from the ground. During terminal stance, the force increased again as the body was accelerated upward prior to the opposite heel strike. Finally, the ground reaction force reduced to zero as the opposite heel contacted the ground and the toe lifted off.

The momentary decrease in ground reaction force that was observed at approximately 105% of the gait cycle was a result of the heel strike impact (Paul, 1976; Watkins & Mathieson, 2009). Slower walking speeds or the use of shoes result in a less pronounced, or non-existent, momentary decrease in the ground reaction force (Watkins & Mathieson, 2009). Participant120 performed all testing barefoot whereas participants 118 and 119 both performed the testing wearing shoes. It was observed from the ground reaction forces of participants 118 and 119

(shown in Sections 6.6.4 and 6.6.5) that their momentary decreases were generally non-existent or very small.

6.6.3 Implant Squeaking and Hip Motion

The observations of the AE and gait analysis of Participant120 showed that squeaking AEs only occurred during the terminal stance stage of the gait cycle. During terminal stance the ground reaction force was increasing from the local minimum to the second of the two peaks (see Figures 6.4b and 6.5b). Figure 6.6 shows the averaged variation of the hip joint forces during slow, normal, and fast walking gaits from a study conducted by Paul (1976). The slow, normal, and fast walking speeds corresponded to participant average linear speeds of 1.10, 1.48, and 2.01 metres per second respectively. The average linear speed of Participant120 during the ten walking tests was 1.19 metres per second. Figure 6.6 shows that, for the study conducted by Paul (1976), the local hip joint force (as opposed to ground reaction force) during terminal stance (30%-50% of the gait cycle) was increasing in the range of approximately 3-4 times body weight. Therefore, during the AEs recorded from the walking tests of Participant120, the implant components were almost certainly subject to high loadings when the squeaking occurred. Furthermore, the results from the testing also showed that the squeaking occurred while the hip joint was in extension. The hip angle was observed to be moving through the range from approximately 5° extension to approximately 15° extension as the squeaking occurred in all tests. Consequently, the squeaking that was observed coincided with a high rate of change of the implant angle. A high rate of change of implant angle indicated a high relative angular velocity between the femoral head and acetabular liner (i.e. the two surfaces of the implant main bearing surface).

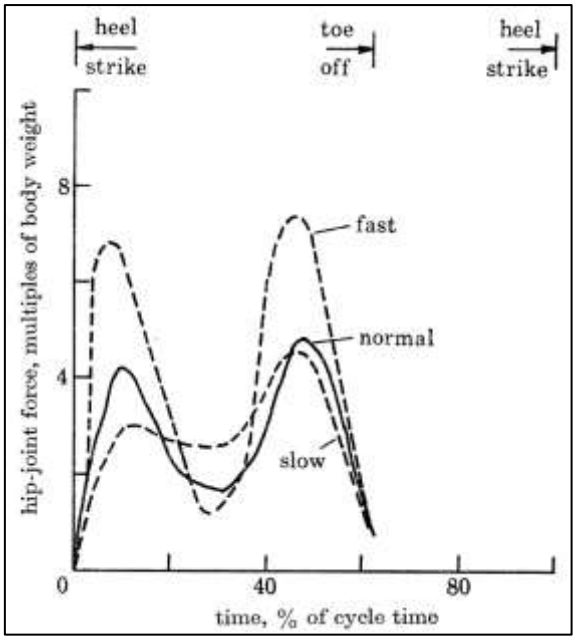

Figure 6.6: Averaged variation during the walking cycle of the resultant hip-joint force at slow, normal, and fast speeds. Reproduced with permission of Royal Society of London, from Paul (1976); permission conveyed through Copyright Clearance Center, Inc.

In addition, it is interesting to note that the squeaking occurred during stages of the gait cycle when the wear rate of the implant surfaces was likely to be increasing to its largest. One of the most commonly used models to quantify the rate of abrasive wear between surfaces is Archard's law (Archard, 1953; Zmitrowicz, 2006). Archard's law has the form:

$$W = k\frac{p_n s}{H} \tag{6.1}$$

Where W the volume of material removed by wear is, k is a dimensionless wear coefficient, p_n is the pressure normal to the contacting surfaces, s is the sliding distance, and H is the hardness of the surface being worn. Therefore, by Equation (6.1), during the gait cycle the main bearing surfaces of the THR undergo the largest amount of wear during the largest normal pressures and largest relative motions between bearing components. The normal pressure of the implant bearing surfaces is correlated with the force through the hip joint. By the plot of Figure 6.6, the joint force during the squeaking of the THR participant would have been increasing rapidly to its maximum force which occurred at approximately 50% of the gait cycle.

Furthermore, from observing the hip angles of the THR participant, the implant bearing surfaces were undergoing large relative rotational displacements during the occurrence of squeaking. Therefore, the squeaking that was observed from Participant120 was likely to be occurring when the THR implant was experiencing large loads, large relative motion, and its largest amount of wear.

6.6.4 Observations from Participant118 – Increased AE Activity

Unlike Participant120, Participant118 did not produce regular squeaking during the combined AE and gait testing. However, there were still some interesting observations from the AE signals of Particpant118. Participant118 produced data for five separate normal walking tests that each contained three consecutive gait cycles. Figures 6.7 and 6.8 show the synchronised AE, ground reaction force, and hip angle data for two separate walking tests of Participant118. It was observed from Figures 6.7 and 6.8 that increased AE activity was consistently present during the stance phase of the gait cycle. Furthermore, the AE signal behaviour that was seen in Figures 6.7 and 6.8 was also observed in the equivalent plots for the three additional walking tests which have not been shown here.

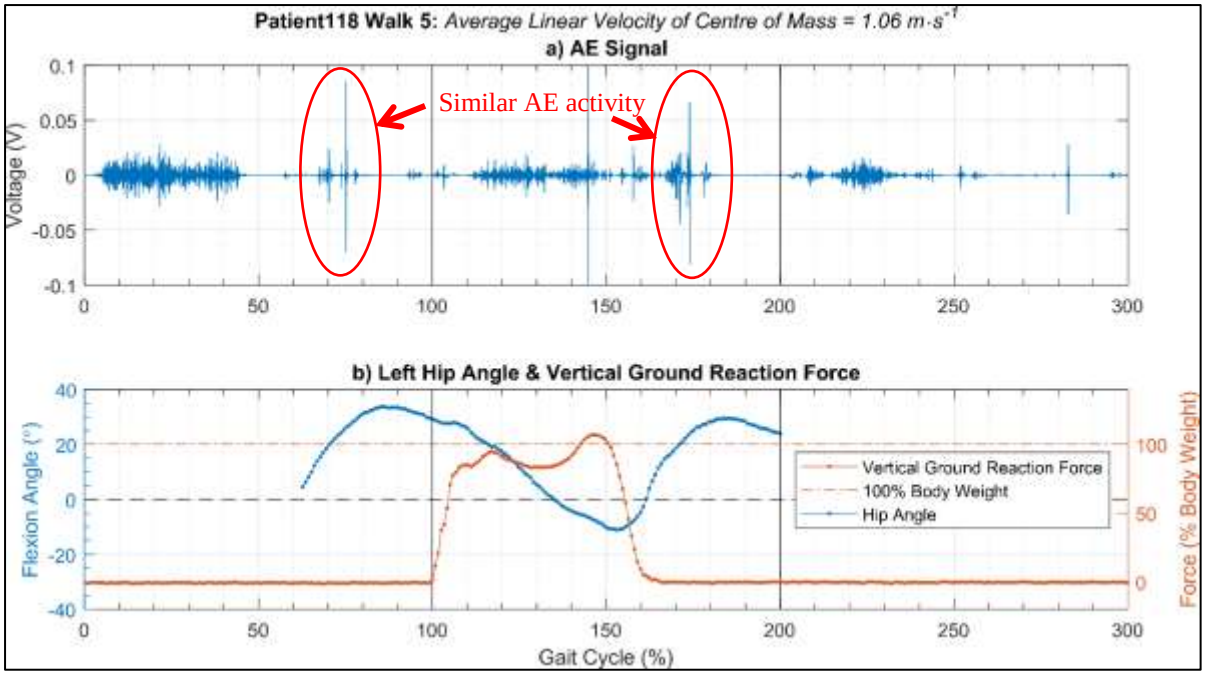

Figure 6.7: a) AE signal from three consecutive gait cycles of a walking test of THR Participant118 with b) synchronised hip angle and ground reaction force data. Note that ground reaction force data was only available for the second consecutive gait cycle.

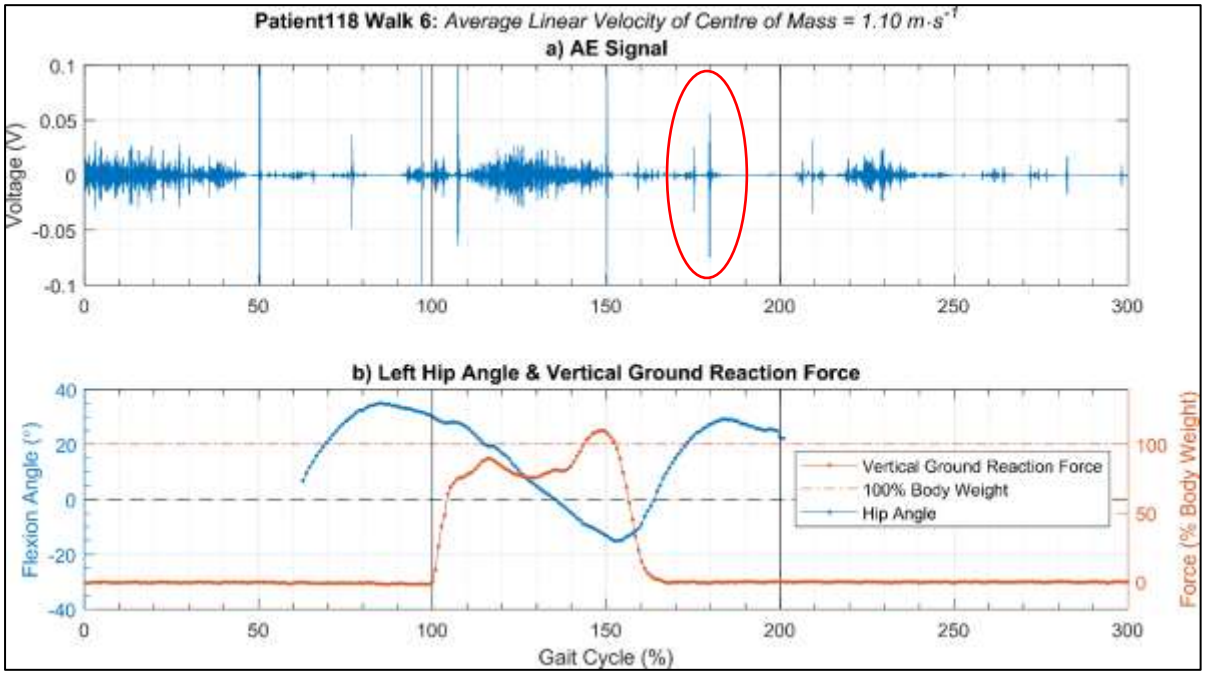

Figure 6.8: a) AE signal from three consecutive gait cycles of a second walking test of THR Participant118 with b) synchronised hip angle and ground reaction force data. Note that ground reaction force data was only available for the second consecutive gait cycle.

The vertical axes of Figures 6.7a and 6.8a indicate that the voltage amplitude of the observed AE activity is considerably lower than what was observed from the squeaks of Figures 6.4a and 6.5a. However, the AE activity that was observed from the walking tests of Participant118 still had voltage amplitudes well above the voltage amplitude of the background noise.

Therefore, the AE activity was almost certainly a result of the motion of the participant. It is possible that some of the low amplitude AE activity could have been the result of external factors such as clothing rubbing over the AE sensor pad or movement of the sensors on the skin surface. However, given that these external factors would almost certainly still be present during the swing phase of the gait cycle, it would be expected that the increased AE activity would also be present during the swing phase. While some AE activity was observed during the swing phases of Figures 6.7a and 6.8a, there were still considerable periods of the swing phases that had no noticeable AE activity.

Another observation from the plots of Figures 6.7 and 6.8 was that there was some AE activity that appeared to be very similar even between the two separate tests on this study participant. For example, see the AE activity indicated by the ellipses in Figures 6.7a and 6.8a. The apparent similarities in the appearance of the AE activity that occurs at the same stage of the gait cycles supports the suggestion that some AEs are likely characteristic of particular implant motions. However, it should be kept in mind that it is also possible the AEs identified by the ellipses in Figures 6.7a and 6.8a are not necessarily caused by the implant components and could be the result of other AE sources (such as other joints in the body, muscle movements, movement or sliding of tendons, fascia, ligaments, and other inherent movement of biological material). Nevertheless, the consistency that has been observed between the AE activity and the walking gait as a result of the combined AE and gait analyses was not previously possible to identify with any certainty.

6.6.5 Observations from Participant119 – Short Duration Impulses

The observations from Participant119 provided further evidence that some AEs from THR patients were likely related to specific motions. Participant119 produced data for five separate

normal walking tests that each contained three consecutive gait cycles. Figure 6.9 and Figure 6.10 show the synchronised AE, ground reaction force, and hip angle data for two separate walking tests of Participant119. It was observed from Figure 6.9 and Figure 6.10 that a lack of AE activity occurred relative to the walking tests of Participant120 and Participant118. However, a number of short duration AE events did occur in Figure 6.9a and Figure 6.10a that had amplitudes comparable to some of the AE events observed from the other participants. Furthermore, some of the AE events that were present in Figure 6.9a and Figure 6.10a showed similarities in appearance and occurred at the same stage of their respective gait cycle. For example, four AE events that appeared to be very similar, in Figure 6.9a and Figure 6.10a, occurred at percentages of the gait cycle that corresponded to heel strike (see annotations in Figure 6.9a and Figure 6.10a). In addition to the AE events that corresponded to heel strike, it was observed that AE events also occurred approximately at the stage of the gait cycle that corresponded to the opposite heel strike. Therefore, the AE activity that was detected from Participant119 showed that some short duration AE events were likely to have been induced by specific motions or gait cycle events such as the impact from heel strikes.

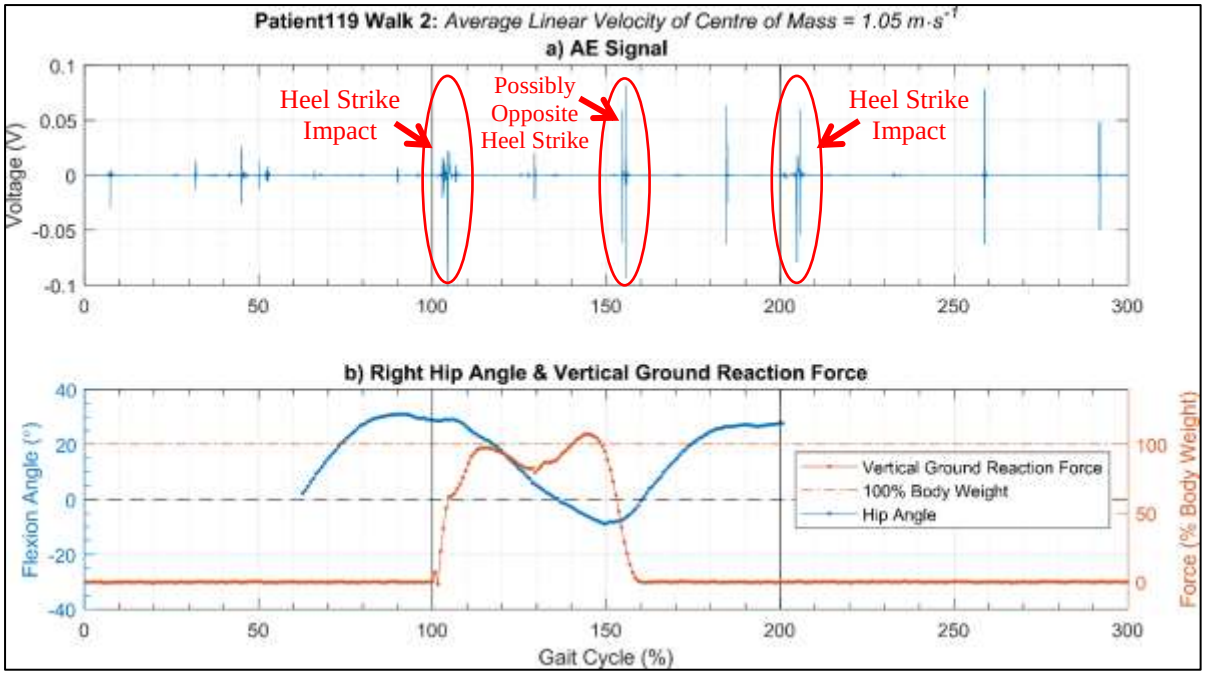

Figure 6.9: a) AE signal from three consecutive gait cycles of a walking test of THR Participant119 with b) synchronised hip angle and ground reaction force data. Note that ground reaction force data was only available for the second consecutive gait cycle.

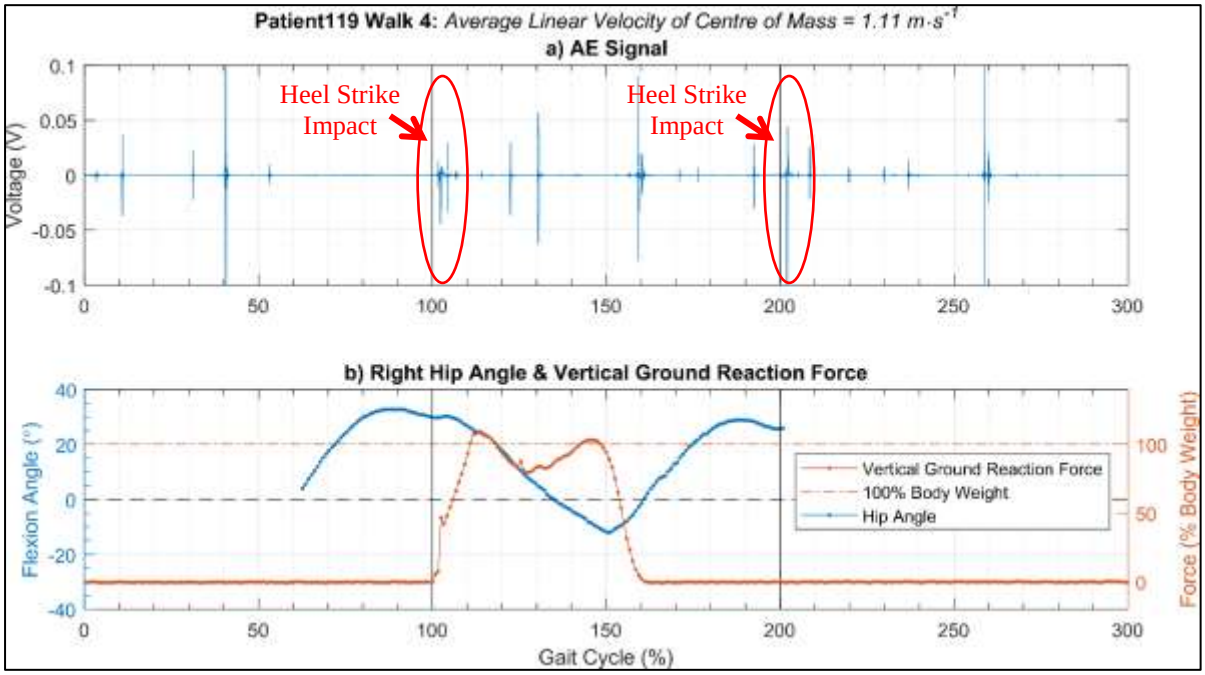

Figure 6.10: a) AE signal from three consecutive gait cycles of a second walking test of THR Participant119 with b) synchronised hip angle and ground reaction force data. Note that ground reaction force data was only available for the second consecutive gait cycle.

6.6.6 Squeak Event Metric Comparison

As noted in Section 6.6.2, all six observed squeaks from Figures 6.4a and 6.5a had a different time domain appearance. These differences in appearance were observed despite the

observation that all the squeaks of the walking tests of Participant120 occurred during similar points within terminal stance. To further investigate the differences in the squeak signals the event detection algorithm was applied to the AE signals of Figures 6.4a and 6.5a. Table 6.7 presents the AE event duration, RMS amplitude, peak amplitude, rise time, and main frequency content for each of the six squeaks of Figures 6.4a and 6.5a. The mean and standard deviation across the six squeaks, for each of the five metrics, is also shown in Table 6.7. It was observed from Table 6.7 that the standard deviation values for all the metrics were seemingly large for all five metrics. Therefore, it was indicated that a large variation was present in the event metrics across the six squeak signals, showing that each individual squeak had unique characteristics.

Table 6.7: Summary of AE event metrics for squeaks of Figures 6.2a and 6.3a.

AE Event		**Duration (ms)**	**RMS Amplitude (V)**	**Peak Amplitude (V)**	**Rise Time (ms)**	**Main Frequency Content (kHz)**
Walk 1 (Figure 6.4)	**Squeak 1**	202	0.37	1.10	126	1.75
	Squeak 2	147	0.20	1.01	117	1.70
	Squeak 3	83	0.08	0.32	50	1.58
Walk 5 (Figure 6.5)	**Squeak 1**	147	0.33	2.54	28	2.37
	Squeak 2	145	0.08	0.40	12	2.44
	Squeak 3	101	0.09	0.33	46	1.46
Mean		**138**	**0.19**	**0.95**	**63**	**1.88**
Standard Deviation		**42**	**0.13**	**0.85**	**47**	**0.42**

The metric values of Table 6.7 were compared to the principal component analysis (PCA) that was performed in Section 5.5.2 to further investigate the variation between the six squeaks. The PCA factor score plot of Figure 5.25b presented the AE events from all study participants that had been revised for a noisy implant against the corresponding first and second principal components (PCs). From Section 5.5.2 it was found that the first principal component (PC1)

represented a linear combination of the event duration, rise time, peak amplitude, and RMS amplitude with each metric having a similar contribution (see Figure 5.26b). Furthermore, it was also found that the second principal component (PC2) represented similar contributions of the same four metrics. However, PC2 was found to contrast the peak and RMS amplitude metrics with the duration and rise time metrics (see Figure 5.26b). The PCA of Section 5.5.2 also found that the third principal component (PC3) represented the main frequency content metric almost completely. Figure 6.11 reproduces the plot of Figure 5.25b and includes the factor scores corresponding to the six squeaking events of Table 6.7 as supplementary observations for comparison (See comment in Section 5.5.2 on supplementary observations).

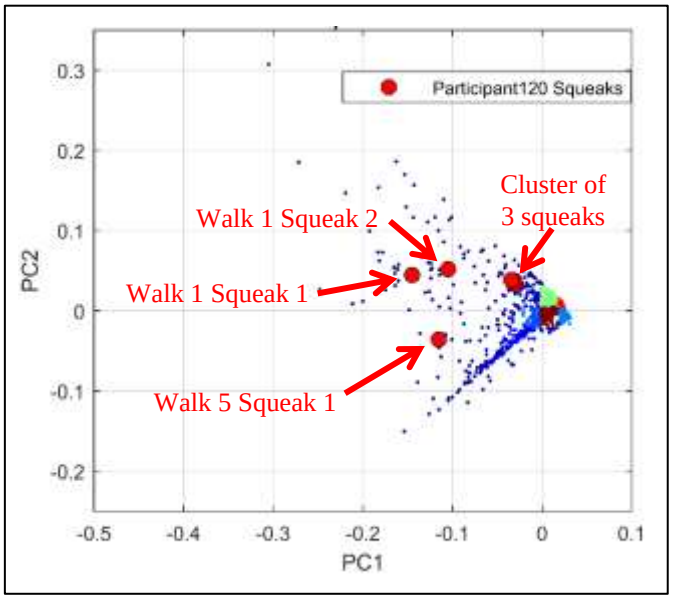

Figure 6.11 Squeaking events from Table 6.7 have been plotted as supplementary PCA points on the PCA factor scores plot from Figure 5.25b which showed the noise revision reason group AE events plotted against the first and second principal components (PC1 and PC2 respectively).

From Figure 6.11 it was observed that three of the squeaks (Squeaks 2 and 3 from Walk 5 and Squeak 3 from Walk 1) were grouped relatively close together. The close grouping of three of the squeaks indicated that overall they had similar values for PC1 and PC2. Consequently, the close grouping indicated that the three squeaks had relatively similar values for the duration, rise time, peak amplitude, and RMS amplitude metrics when considering the overall variation. The remaining three squeaks were observed to have considerably more negative PC1 values.

However, two of the squeaks (Squeaks 1 and 2 from Walk 1) had PC2 values that were similar to the PC2 values of the other three squeaks. Overall, Figure 6.11 showed that three of the squeaks (Squeaks 1 and 2 from Walk 1 and Squeak 1 from Walk 5) had characteristics that were commonly seen in THR participants that were revised for noisy implants. The characteristics of the other three squeaks (Squeaks 2 and 3 from Walk 5 and Squeak 3 from Walk 1) had locations in Figure 6.11 that (from the conclusions of Section 5.5.2) did not necessarily indicate which revision group of THR participants the AE events could be categorised into.

Since the squeaking events were observed to have occurred at similar stages of the gait cycle it was likely that the squeaks were induced by the same underlying mechanisms. However, the differences observed from the metrics of the six squeaking events suggest that not all squeaks that occur at the same stage of the gait cycle have similar signal properties. Nonetheless, upon further investigation, it was found that the frequency content of the squeaks exhibited similar trends when considering the temporal behaviour of the frequency content of the squeaks. Therefore, while the metrics of duration, rise time, peak amplitude, and RMS amplitude may have had a sizeable variance across the squeaks, the frequency responses of the AEs were generally comparable.

The implication of the differing metrics from seemingly similar AEs is that the current event detection method and subsequent event metric calculations may be limited in their ability to make accurate comparisons of different AE events. However, analysing AE signals according to phases of the gait cycle has now been made possible and may help to provide clearer discrimination between AEs from different participant groups. By restricting analysis of AE signals to specific regions of the gait cycle, better comparisons of the AE parameters are able to be made. For example, the metrics of AE events that belong to the subset of events that

occurred between 30%-50% of the gait cycle could be compared. As a result, any similarities in the AE event parameters may become more apparent since potentially unrelated AE events would be ignored. However, a much larger set of combined AE and gait tracking data (beyond the three study participants tested to date) is required before this additional analysis can be undertaken.

Furthermore, the event detection algorithm could potentially be disregarded altogether in favour of comparing the temporal signal properties of AE signals in particular regions of the gait cycle. The thresholding values used in the event detection process were arbitrary. Consequently, eliminating the need for an event detection process would remove an element of subjectivity from the overall AE analysis. Additionally, inter-participant comparisons of AE signals could be made more reliable using comparisons limited to gait cycle stages. Given that each participant performs motions at self-selected speeds, AEs from motions performed at different rates could now be reliably compared by matching the dynamic loading and orientation profiles of implant components.

6.7 Summary

A method for detecting THR AEs while simultaneously collecting gait data has been developed. The existing AE monitoring hardware and data collection protocols of Chapter 3 were employed to collect the AE data. In conjunction with the AE data, protocols to collect gait data using a BTS Bioengineering motion analysis system have been developed. Furthermore, the gait data collection protocol incorporates the use of three IMUs to collect inertial data from the lower limb. The inertial data collection will contribute to ongoing and future research with the aim of allowing the gait analysis to be conducted solely from IMU data. The developed methodology for the combined analysis of AE and gait data has

successfully been used to collect simultaneous AE and gait data from three THR participants. Methodology has been developed to apply the gait data to a computer musculoskeletal model, through OpenSim software, to estimate hip joint motion. The computer modelling process ultimately yields joint angle estimates, for all joints of the torso and lower limbs, which are synchronised in time with the AE data. Furthermore, the OpenSim software has the capability of estimating the corresponding forces at the joints by factoring in contributions from the muscles and force plate data.

The initial results from the combined AE and gait analysis of the three THR participants who underwent the testing were promising. Substantial AEs have been observed to occur only during the stance phase of the gait cycle in two out of three participants. Repeatable audible squeaking was observed from one of these participants during terminal stance (i.e. 30%-50% of the gait cycle). The squeaking was observed to occur from approximately 5°-15° extension when high joint loading and rapid changes in hip angle were taking place, which is also likely to be the conditions which lead to the largest wear rate of the main bearing surfaces of the THR implant. Participant 119 was observed to have less AE activity relative to the other participants. However, this participant had large amplitude, short duration AEs that consistently occurred and seemed to correspond with the heel strike events of the gait cycle.

The research objectives of Section 1.5 posed the question: Can any correlations be seen between implant joint angles and audible AEs? The results from the initial dataset of three participants has shown that correlations between audible squeaking and hip joint angles do seem to exist. Furthermore, correlations seemed to also exist between overall increased AE activity and the load bearing phases of the walking gait cycle. A further question posed in Section 1.5 asked if any observed correlations were able to improve the ability of the AE technique to categorise AEs into revision reason groups. The answer to this second question is

less clear from the results of the three THR participants who underwent combined AE monitoring and gait analysis. From the results of the three participants who underwent the combined analysis, it was evident that the participant who produced audible squeaking would have been categorised into the noise revision group as a result of the presence of squeaks. Alternatively, this same categorisation could have been reached without considering the gait data. However, the combined technique has shown that being able to compare AE signals by restricting them to specific stages of the gait cycle will likely result in clearer identification of AEs with similar characteristics. Nevertheless, further data is required before meaningful conclusions can be drawn that relate AEs with stages of the gait cycle and associated implant loading. Since the groundwork has now been laid for further THR participants to be tested with the combined AE monitoring and gait analysis, it is anticipated that ongoing research will verify the observations presented in this chapter and continue to show promise of the diagnostic potential of the AE monitoring technique.

Chapter 7 Conclusions & Future Work

7.1 Conclusions

This thesis presented an analysis of the AE monitoring technique as a potential diagnostic tool for the assessment and early diagnosis of THR implant complications. The research hypothesis was that THR implant complications would have unique and identifiable AE characteristics that could be used to assist orthopaedic surgeons in making more informed and prompt implant complication diagnoses. Analysis of the AE data showed promise that the AE monitoring technique could be a potential diagnostic tool. Some distinctions were found in the distributions of AE event metric values between different sets of study participants grouped by implant revision reasons. However, the AE data analysis was not able to conclusively show that AEs from THR participants could be reliably categorised by their characteristics and be associated with distinct implant revision reasons.

A procedure was also developed to collect AE data of THR participants concurrently with gait tracking data. Data was collected from three THR study participants using the combined AE and gait tracking methodology. Initial results from the combined testing found that some AEs occurred repeatedly at specific stages of the gait cycle. In particular, audible squeaking was observed to occur consistently during the terminal stance phase of the gait cycle. Consequently, the preliminary results of the combined AE monitoring and concurrent gait analysis strengthened the evidence supporting the AE monitoring technique as a potential diagnostic tool of implant complications. Nevertheless, more data is required to verify the outcomes from combined AE monitoring and gait analysis.

An initial analysis on the *in-vivo* and *in-vitro* AE signals that investigated time and frequency domain behaviour of the AEs showed that the majority of the AE activity in the *in-vivo* signals was from short duration (<20 ms) AE events. However, a number of long duration AE events were observed from THR participants that were generally associated with audible squeaking. The *in-vivo* non-squeaking AEs (from both the THR participants and the control participants) were generally found to have main frequency content in the 1-4 kHz range. The frequency responses of AE signals containing squeaking typically had main frequency content in the same the range of 1-4 kHz and could exhibit distinct narrow-band peaks or more broad-band noisy peaks. Additionally, many of the squeaking AEs were found to exhibit strong harmonic signals and transient frequency content.

In-vitro testing on implant components investigated the AEs from the interaction of the main bearing interface in addition to the interaction of the femoral components at the trunnion Morse taper (TMT) interface. The *in-vitro* main bearing interaction AE signals primarily contain long duration squeaking AE events. Consequently, it is likely that a substantial amount of the short duration AE events detected *in-vivo* are a result of interactions between other components of the overall joint system (such as muscle, soft tissue, bone, ligaments, fascia, tendons, bone to implant interactions, and possibly even indirect contributions from the knee joint). Furthermore, the *in-vitro* main bearing interface AEs had frequency content in the same range as the *in-vivo* AEs. Conversely, the *in-vitro* AEs from the TMT interactions had frequency content upwards of 12 kHz, with the primary content typically in the 18-22 kHz range. Consequently, it was concluded that squeaking observed from THR participants *in-vivo* was much more likely to have been induced by interactions of the main bearing interface than having been caused directly by interactions at the TMT.

An event detection algorithm was developed to automate the detection of regions of interest from the large dataset of AE signals. The event detection algorithm performed well at identifying THR AE events and provided good distinction between events that occurred closely in time. Event metrics of duration, rise time, RMS amplitude, peak amplitude, and main frequency content was determined for each AE event. The AE events were then grouped using the control participants and the clinical revision reasons. Results from the statistical analysis indicated that the single metrics alone were not sufficient to categorise participant AE data. Therefore, two separate multidimensional metric analyses was carried out, in the form of two-dimensional scatter plot comparisons and a principal component analysis (PCA), to investigate if combining multiple metrics would result in more distinct differences between revision reason groupings.

The two-dimensional event metric analysis showed that it may be possible to categorise a THR AE event into the noise revision group but it was more likely that a random AE event would have event metrics that are common to all participant groupings. However, correlations between metrics were observed that were unable to be noticed from the statistical distribution analysis. The PCA showed that a definite difference in the distributions of different groups of AE events was apparent, particularly contrasting the loosening group of THR participants. The PCA result suggests it is likely that particular AE events may be more (or less) prevalent in THR patients with particular implant complications such as loosening. The PCA results offered the strongest support of the potential diagnostic capability of the AE monitoring technique. However, further research is needed to identify and examine the AE event which had the most influence on the differing PCA results as those events are likely to be characteristic to particular implant failure modes.

Overall, the initial results from the AE analysis show some promise in the diagnostic potential of the AE monitoring technique. However, the results fell well short of being able to lead to a reliable failure diagnosis in their current state. Consequently, a system was developed to collect concurrent AE, gait, IMU, and force plate data to allow the synchronisation of AE data with specific motions and loadings of the implant. Three THR study participants have been tested using the combined AE and gait analysis system. The initial data from the combined system has shown promising intra- and inter-participant consistency of the AE and gait analysis results. Substantial AE activity was observed to occur repeatedly during the stance phase of the gait cycle. Furthermore, audible squeaking was observed to occur consistently during the terminal stance phase of one participant. Consequently, the occurrence of the audible squeaking was shown to correspond well with large vertical ground reaction forces and high rates of change of hip joint angle.

This study has indicated that the AE device shows promise as a potential diagnostic tool for assessing the condition of THR implants. While a large amount of AE data has been shown to be common across the control and various implant revision reason participant groups, results have shown that some noticeably different AE characteristics did exist between participant groups. Initial testing from combined AE and gait analysis has indicated relationships likely exist between AEs and particular implant motions and loadings. However, more research is needed to verify the findings of this study and provide more conclusive evidence regarding the diagnostic capability of the AE monitoring technique.

7.2 Future Work

A number of specific areas of future research related to this thesis have been identified and are presented in this section. The most obvious and important next step of this study is to collect

more data from THR patients using the combined AE monitoring and gait analysis. The developed sensing system with AE and gait tracking, along with a detailed data collection protocol has been developed and is now available to allow the combined AE and gait data to be collected easily. Therefore, recruitment of more study participants who have THR implants is required. Due to the limited availability of suitable THR patients within the Canterbury region, an average of one or two study participants are usually recruited per month. Consequently, the collection of data from large amount of study participants happens over a large period of time. However, since detailed data collection protocols for the combined AE monitoring and gait analysis have now been developed in the work of this thesis, it would be beneficial to direct more focus on the recruitment of suitable study participants and possibility of expanding the region that participants are recruited from.

A major limitation of the *in-vitro* testing that was conducted was the use of manual implant manipulations for the trunnion Morse taper (TMT) rotations. The nature of the manual manipulation meant that the force, speed, and rotation that was applied to manipulate the implant components was not able to be controlled as accurately as with the flexion-extension manipulations that used robotic control. Therefore, it would be desirable to conduct further TMT rotation testing using robotic manipulation. The main challenge with performing these tests robotically is finding an effective method to grip the femoral head and femoral stem in order to apply the desired rotations.

The analysis of the *in-vitro* data of this study found that there was consistency in the occurrence of some sequential AE events. For example, some *in-vitro* audible squeaking AE events were found to be followed by a short duration AE event with a consistent timing gap between the events. Since the study now has the ability to synchronise *in-vivo* AE signals with the motion of the hip joint, it would be interesting to investigate the occurrences of sequential AE events

from the *in-vivo* environment. This additional analysis will be enabled as more data is collected using the combined AE-gait analysis system, where AEs can be linked to temporal gait motion, limb angles and implant loads.

Another important aspect of the AE analysis presented in this thesis was that the event detection method that was applied to the data incorporated arbitrarily chosen voltage thresholds for event detection. While these voltage and time threshold values were iteratively tested and manually optimised, there is no indication that these are necessarily optimal values. Therefore, it would be desirable to perform a more comprehensive study that investigates the effect of different threshold values used in the event detection thresholds on the subsequent event metric analysis. Additionally, the event metric analysis that was performed in this thesis was developed in such a way that it can somewhat easily be applied to AE events from any alternative AE event detection methods. Furthermore, since the combined AE monitoring and gait analysis allows AEs to be temporally matched to the gait cycle, a gait based event detection or signal analysis could be implemented as an alternative to methods that incorporate arbitrary voltage thresholds. For example, once a larger dataset with gait analysis is developed, the event metric comparisons could be re-created, where only events within the terminal stance phase of gait could be included (or any other subset of the gait cycle could be included or excluded). Such gait-stage selection of a subset of data has the potential to remove AEs unrelated to implant mechanics and could potentially lead to improved distinction between revision reason groups.

In addition, a useful improvement of the current AE monitoring system would be to develop an on-patient or wireless data logging method. Currently, the AE sensors are required to be tethered to a laptop computer for the data acquisition. The tethering of the AE sensors limits the distance the study participants can move away from the data acquisition laptop. Consequently, motions like walking can only be performed in a limited space. Furthermore,

the cables from the AE sensors occasionally obstruct the line-of-sight of the motion capture cameras to the reflective markers positioned on the study participant. This obstruction introduces error to the marker positions and occasionally the marker positions cannot be tracked for a small number of data frames. The development of an on-patient or wireless AE and IMU data collection method would allow a more streamlined testing procedure. One of the major challenges with developing such a system is that the high sampling rate of the AE data (100 kHz) generates large amounts of data in a short time span. Therefore, any wireless or on-patient system must be able to reliably and quickly transfer or store the data during the testing. From observing the results of this thesis, it would be possible to reduce the sampling rate of the AE sensors and still capture important AEs. However, given that some AE content was detected up to 50 kHz, a high sample rate is desirable for the AE monitoring. In the longer term, it would be desirable to have a system that could be incorporated into a wearable device, where patients could undertake AE monitoring over many days during normal life, rather than for a short period in a clinical environment. However, significant additional research is needed before this could be undertaken.

Furthermore, a key part of future research would be to develop and implement population based inverse kinematics models that predict the gait based on the IMU data only. The current combined AE and gait system relies on access to a motion capture facility with trained personnel to operate the system. These requirements limit the ability for such systems to be used in clinical practice in regional settings, restricting access to the technology. The development of a system to track gait based only off the IMUs would allow for a much more portable, cheaper, and easily accessible combined AE monitoring and gait analysis to be applied to THR patients.

A natural extension to the current gait analysis is to perform inverse dynamics using the computer musculoskeletal model to estimate the muscle reaction forces of the lower limbs and the associated joint loads. The ability to directly compare AE signals and hip joint loadings would provide a much greater insight into *in-vivo* implant biomechanics and the relationship between implant AEs and implant loads. In particular, an investigation of the relationship between implant squeaking frequencies and implant loading would be of great interest. *In-vitro* studies from literature have suggested a correlation exists between implant load and squeaking frequency. While transient squeaking frequencies have been seen *in-vivo* within this research, that data was not linked to any gait tracking, so it is not possible to retrospectively determine whther those transient squeaking frequencies are linked to transient implant loads. The combined AE monitoring and gait analysis technique of this thesis, especially if combined with inverse dynamic modelling, will provide the ability to examine if any such correlation exist within the *in-vivo* environment.

Finally, the event detection and subsequent event metric comparison methods presented within this thesis are only one approach to data analysis. While they make sense as an initial approach to understand the characteristics of the AE data, the use of machine learning, neural networks and genetic algorithms has the potiental to identify other traits that exist within the datset. While such additional analyses could be very complicated and prolonged, and may require a much larger dataset than is currently available, there is significant potential for new insight to be gained from this approach.

References

Abdi, H., & Williams, L. J. (2010). Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics, 2*(4), 433-459. doi: 10.1002/wics.101

Abernethy, A., Hanrahan, S., Kippers, V., Mackinnon, L., & Pandy, M. (2005). The biophysical foundations of human movement.

Abu-Amer, Y., Darwech, I., & Clohisy, J. C. (2007). Aseptic loosening of total joint replacements: mechanisms underlying osteolysis and potential therapies. *Arthritis Res Ther, 9 Suppl 1*(Suppl 1), S6. doi: 10.1186/ar2170

Ahadi, M., & Bakhtiar, M. S. (2010). Leak detection in water-filled plastic pipes through the application of tuned wavelet transforms to Acoustic Emission signals. *Applied Acoustics, 71*(7), 634-639. doi: 10.1016/j.apacoust.2010.02.006

Akay, A. (2002). Acoustics of friction. *J Acoust Soc Am, 111*(4), 1525-1548. doi: 10.1121/1.1456514

Any questions? (1969). *British Medical Journal, 4*(5679), 351-351.

Archard, J. F. (1953). Contact and Rubbing of Flat Surfaces. *Journal of Applied Physics, 24*(8), 981-988. doi: 10.1063/1.1721448

Australian Orthopaedic Association National Joint Replacement Registry. (2015). Annual Report. Retrieved from: https://aoanjrr.sahmri.com/annual-reports-2015

Baranov, V., Kudryavtsev, E., Sarychev, G., & Schavelin, V. (2007). Chapter 1 Friction of solids and nature of acoustic emission *Acoustic Emission in Friction - Tribology and Interface Engineering Series* (Vol. 53, pp. 1-36): Elsevier.

Beaulieu, M. L., Lamontagne, M., & Beaule, P. E. (2010). Lower limb biomechanics during gait do not return to normal following total hip arthroplasty. *Gait Posture, 32*(2), 269-273. doi: 10.1016/j.gaitpost.2010.05.007

Bennett, D., Humphreys, L., O'Brien, S., Kelly, C., Orr, J. F., & Beverland, D. E. (2008). Gait kinematics of age-stratified hip replacement patients--a large scale, long-term follow-up study. *Gait Posture, 28*(2), 194-200. doi: 10.1016/j.gaitpost.2007.11.010

Blausen.com staff. (2014). Medical gallery of Blausen Medical 2014. *WikiJournal of Medicine, 1*(2). doi: 10.15347/wjm/2014.010

Bogert, B., Healy, M., & Tukey, J. (1963). *The quefrency analysis of time series for echoes: cepstrum, pseudo-autocovariance, cross-cepstrum and saphe craking*. Paper presented at the Proc. Symp. On Time Series Analysis.

Brinckmann, P., Frobin, W., & Leivseth, G. (2002). *Musculoskeletal biomechanics*. Stuttgart: Thieme.

Brockett, C. L., Williams, S., Jin, Z., Isaac, G. H., & Fisher, J. (2013). Squeaking hip arthroplasties: a tribological phenomenon. *J Arthroplasty, 28*(1), 90-97. doi: 10.1016/j.arth.2012.01.023

Browne, M., Jeffers, J. R., & Saffari, N. (2010). Nondestructive evaluation of bone cement and bone cement/metal interface failure. *J Biomed Mater Res B Appl Biomater, 92*(2), 420-429. doi: 10.1002/jbm.b.31530

Browne, M., Roques, A., & Taylor, A. (2005). The acoustic emission technique in orthopaedics - a review. *Journal of Strain Analysis for Engineering Design, 40*(1), 59-79. doi: 10.1243/030932405x7638

Burge, A. J. (2015). Total hip arthroplasty: MR imaging of complications unrelated to metal wear. *Semin Musculoskelet Radiol, 19*(1), 31-39. doi: 10.1055/s-0034-1396765

Canadian Institute for Health Information. (2015). *Hip and Knee Replacements in Canada: Canadian Joint Replacement Registry 2015 Annual Report* Retrieved from https://www.cihi.ca/en/types-of-care/specialized-services/joint-replacements

Cappozzo, A. (1984). Gait Analysis Methodology. *Human Movement Science, 3*(1-2), 27-50. doi: 10.1016/0167-9457(84)90004-6

Cappozzo, A., Catani, F., Croce, U. D., & Leardini, A. (1995). Position and orientation in space of bones during movement: anatomical frame definition and determination. *Clin Biomech (Bristol, Avon), 10*(4), 171-178. doi: 10.1016/0268-0033(95)91394-t

Castellanos, J., & Axelrod, D. (1990). Effect of habitual knuckle cracking on hand function. *Ann Rheum Dis, 49*(5), 308-309. doi: 10.1136/ard.49.5.308

Chandran Suja, V., & Barakat, A. I. (2018). A Mathematical Model for the Sounds Produced by Knuckle Cracking. *Sci Rep, 8*(1), 4600. doi: 10.1038/s41598-018-22664-4

Chevillotte, C., Trousdale, R. T., Chen, Q., Guyen, O., & An, K. N. (2010). The 2009 Frank Stinchfield Award: "Hip squeaking": a biomechanical study of ceramic-on-ceramic bearing surfaces. *Clinical Orthopaedics and Related Research®, 468*(2), 345-350. doi: 10.1007/s11999-009-0911-x

Chung, M. J., & Wang, M. J. (2010). The change of gait parameters during walking at different percentage of preferred walking speed for healthy adults aged 20-60 years. *Gait Posture, 31*(1), 131-135. doi: 10.1016/j.gaitpost.2009.09.013

Clohisy, J. C., Carlisle, J. C., Trousdale, R., Kim, Y. J., Beaule, P. E., Morgan, P., . . . Millis, M. (2009). Radiographic evaluation of the hip has limited reliability. *Clinical Orthopaedics and Related Research®, 467*(3), 666-675. doi: 10.1007/s11999-008-0626-4

Currier, J. H., Anderson, D. E., & Van Citters, D. W. (2010). A proposed mechanism for squeaking of ceramic-on-ceramic hips. *Wear, 269*(11-12), 782-789. doi: 10.1016/j.wear.2010.08.006

Dattani, R. (2007). Femoral osteolysis following total hip replacement. *Postgrad Med J, 83*(979), 312-316. doi: 10.1136/pgmj.2006.053215

David, L., & Dumitrascu, D. L. (2017). The bicentennial of the stethoscope: a reappraisal. *Clujul Med, 90*(3), 361-363. doi: 10.15386/cjmed-821

Davies, J. P., Tse, M. K., & Harris, W. H. (1996). Monitoring the integrity of the cement-metal interface of total joint components in vitro using acoustic emission and ultrasound. *Journal of Arthroplasty, 11*(5), 594-601. doi: 10.1016/S0883-5403(96)80115-X

Davis, R. B., Ounpuu, S., Tyburski, D., & Gage, J. R. (1991). A Gait Analysis Data-Collection and Reduction Technique. *Human Movement Science, 10*(5), 575-587. doi: 10.1016/0167-9457(91)90046-Z

Delaunay, C., Hamadouche, M., Girard, J., Duhamel, A., & So, F. G. (2013). What are the causes for failures of primary hip arthroplasties in France? *Clinical Orthopaedics and Related Research®, 471*(12), 3863-3869. doi: 10.1007/s11999-013-2935-5

Delp, S. L., Anderson, F. C., Arnold, A. S., Loan, P., Habib, A., John, C. T., . . . Thelen, D. G. (2007). OpenSim: open-source software to create and analyze dynamic simulations of movement. *IEEE Trans Biomed Eng, 54*(11), 1940-1950. doi: 10.1109/TBME.2007.901024

Drees, P., Eckardt, A., Gay, R. E., Gay, S., & Huber, L. C. (2007). Mechanisms of disease: Molecular insights into aseptic loosening of orthopedic implants. *Nat Clin Pract Rheumatol, 3*(3), 165-171. doi: 10.1038/ncprheum0428

Ecker, T. M., Robbins, C., van Flandern, G., Patch, D., Steppacher, S. D., Bierbaum, B., & Murphy, S. B. (2008). Squeaking in total hip replacement: no cause for concern. *Orthopedics, 31*(9), 875-876, 884. doi: 10.3928/01477447-20080901-11

Eitzen, D. G., & Wadley, H. N. G. (1984). Acoustic-Emission - Establishing the Fundamentals. *Journal of research of the National Bureau of Standards, 89*(1), 75-100. doi: 10.6028/jres.089.008

FitzPatrick, A. J., Rodgers, G. W., Hooper, G. J., & Woodfield, T. B. F. (2017a). *Cepstrum Analysis for Determining the Fundamental Frequency of Total Hip Replacement Acoustic Emissions*. Paper presented at the International Federation of Automatic Control (IFAC) World Congress, Toulouse, France. doi:10.1016/j.ifacol.2017.08.1642

FitzPatrick, A. J., Rodgers, G. W., Hooper, G. J., & Woodfield, T. B. F. (2017b). Development and validation of an acoustic emission device to measure wear in total hip replacements in-vitro and in-vivo. *Biomedical Signal Processing and Control, 33*, 281-288. doi: 10.1016/j.bspc.2016.12.011

Gao, X. J., Murota, K., Tomita, Y., Ono, M., Higo, Y., & Nunomura, S. (1990). Evaluation of the Fixation of Artificial Hip-Joint by Acoustic-Emission. *Japanese Journal Of Applied Physics Part 1-Regular Papers Short Notes & Review Papers, 29*(S1), 215-217. doi: 10.7567/Jjaps.29s1.215

Garellick, G., Kärrholm, J., Lindahl, H., Malchau, H., Rogmark, C., & Rolfson, O. (2015). Swedish Hip Arthroplasty Register: Annual Report 2014. Retrieved from: http://www.shpr.se/en/Publications/DocumentsReports.aspx

Gholizadeh, S., Leman, Z., & Baharudin, B. T. H. T. (2015). A review of the application of acoustic emission technique in engineering. *Structural Engineering and Mechanics, 54*(6), 1075-1095. doi: 10.12989/sem.2015.54.6.1075

Glaser, D. (2008). *Development and Implementation of Mathematical Modeling, Vibration and Acoustic Emission Technique to Correlate In Vivo Kinematics, Kinetics and Sound in Total Hip Arthroplasty with Different Bearing Surfaces.* (PhD), University of Tennessee. Retrieved from http://trace.tennessee.edu/utk_graddiss/437

Glaser, D., Komistek, R. D., Cates, H. E., & Mahfouz, M. R. (2008). Clicking and squeaking: in vivo correlation of sound and separation for different bearing surfaces. *J Bone Joint Surg Am, 90 Suppl 4*(Supplement 4), 112-120. doi: 10.2106/JBJS.H.00627

Glaser, D., Komistek, R. D., Cates, H. E., & Mahfouz, M. R. (2010). A non-invasive acoustic and vibration analysis technique for evaluation of hip joint conditions. *J Biomech, 43*(3), 426-432. doi: 10.1016/j.jbiomech.2009.10.005

Glyn-Jones, S., Palmer, A. J., Agricola, R., Price, A. J., Vincent, T. L., Weinans, H., & Carr, A. J. (2015). Osteoarthritis. *Lancet, 386*(9991), 376-387. doi: 10.1016/S0140-6736(14)60802-3

Goldberg, J. R., Gilbert, J. L., Jacobs, J. J., Bauer, T. W., Paprosky, W., & Leurgans, S. (2002). A multicenter retrieval study of the taper interfaces of modular hip prostheses. *Clinical Orthopaedics and Related Research*®(401), 149-161. doi: 10.1097/00003086-200208000-00018

Grzeszick, R., Plinge, A., & Fink, G. A. (2017). Bag-of-Features Methods for Acoustic Event Detection and Classification. *Ieee-Acm Transactions on Audio Speech and Language Processing, 25*(6), 1242-1252. doi: 10.1109/Taslp.2017.2690574

Guillin, R., Marchand, A. J., Roux, A., Niederberger, E., & Duvauferrier, R. (2012). Imaging of snapping phenomena. *Br J Radiol, 85*(1018), 1343-1353. doi: 10.1259/bjr/52009417

Hallab, N. J., Messina, C., Skipor, A., & Jacobs, J. J. (2004). Differences in the fretting corrosion of metal-metal and ceramic-metal modular junctions of total hip replacements. *J Orthop Res, 22*(2), 250-259. doi: 10.1016/S0736-0266(03)00186-4

Hamner, S. R., Seth, A., & Delp, S. L. (2010). Muscle contributions to propulsion and support during running. *J Biomech, 43*(14), 2709-2716. doi: 10.1016/j.jbiomech.2010.06.025

Hannouche, D., Delambre, J., Zadegan, F., Sedel, L., & Nizard, R. (2010). Is there a risk in placing a ceramic head on a previously implanted trunion? *Clinical Orthopaedics and Related Research®, 468*(12), 3322-3327. doi: 10.1007/s11999-010-1505-3

Harris, G. F., & Wertsch, J. J. (1994). Procedures for gait analysis. *Arch Phys Med Rehabil, 75*(2), 216-225. doi: 10.5555/uri:pii:0003999394903999

Hayter, C. L., Koff, M. F., & Potter, H. G. (2012). Magnetic resonance imaging of the postoperative hip. *J Magn Reson Imaging, 35*(5), 1013-1025. doi: 10.1002/jmri.23523

He, Y. Y., Zhang, X. M., & Friswell, M. I. (2010). Observation of time-frequency characteristics of the acoustic emission from defects in rolling element bearings. *Insight, 52*(8), 412-418. doi: 10.1784/insi.2010.52.8.412

Hooper, G. (2013). The ageing population and the increasing demand for joint replacement. *N Z Med J, 126*(1377), 5-6.

Hooper, G., Lee, A. J., Rothwell, A., & Frampton, C. (2014). Current trends and projections in the utilisation rates of hip and knee replacement in New Zealand from 2001 to 2026. *N Z Med J, 127*(1401), 82-93.

Hothan, A., Huber, G., Weiss, C., Hoffmann, N., & Morlock, M. (2011). The influence of component design, bearing clearance and axial load on the squeaking characteristics of ceramic hip articulations. *J Biomech, 44*(5), 837-841. doi: 10.1016/j.jbiomech.2010.12.012

Hutchings, I. M. (2003). *Friction, lubrication, and wear of artificial joints*. Bury St. Edmunds, U.K.: Professional Engineering Pub.

Imbuldeniya, A. M., Pearce, S. J., Walter, W. L., Zicat, B. A., & Walter, W. K. (2013). Squeaking: Current knowledge and how to avoid it. *Curr Rev Musculoskelet Med, 6*(4), 342-349. doi: 10.1007/s12178-013-9181-z

Iosa, M., Fusco, A., Marchetti, F., Morone, G., Caltagirone, C., Paolucci, S., & Peppe, A. (2013). The golden ratio of gait harmony: repetitive proportions of repetitive gait phases. *Biomed Res Int, 2013*(3), 918642. doi: 10.1155/2013/918642

Jacobs, J. J., Cooper, H. J., Urban, R. M., Wixson, R. L., & Della Valle, C. J. (2014). What do we know about taper corrosion in total hip arthroplasty? *J Arthroplasty, 29*(4), 668-669. doi: 10.1016/j.arth.2014.02.014

Jarrett, C. A., Ranawat, A. S., Bruzzone, M., Blum, Y. C., Rodriguez, J. A., & Ranawat, C. S. (2009). The squeaking hip: a phenomenon of ceramic-on-ceramic total hip arthroplasty. *J Bone Joint Surg Am, 91*(6), 1344-1349. doi: 10.2106/JBJS.F.00970

Kadaba, M. P., Ramakrishnan, H. K., & Wootten, M. E. (1990). Measurement of lower extremity kinematics during level walking. *J Orthop Res, 8*(3), 383-392. doi: 10.1002/jor.1100080310

Kadambe, S., & Boudreauxbartels, G. F. (1992). Application of the Wavelet Transform for Pitch Detection of Speech Signals. *IEEE Transactions on Information Theory, 38*(2), 917-924. doi: Doi 10.1109/18.119752

Kaniusas, E. (2015). Sensing by Acoustic Biosignals. In E. Kaniusas (Ed.), *Biomedical Signals and Sensors II: Linking Acoustic and Optic Biosignals and Biomedical Sensors* (pp. 1-90). Berlin, Heidelberg: Springer Berlin Heidelberg.

Kawchuk, G. N., Fryer, J., Jaremko, J. L., Zeng, H., Rowe, L., & Thompson, R. (2015). Real-time visualization of joint cavitation. *PLoS One, 10*(4), e0119470. doi: 10.1371/journal.pone.0119470

Keurentjes, J. C., Kuipers, R. M., Wever, D. J., & Schreurs, B. W. (2008). High incidence of squeaking in THAs with alumina ceramic-on-ceramic bearings. *Clinical Orthopaedics and Related Research®, 466*(6), 1438-1443. doi: 10.1007/s11999-008-0177-8

Khan-Edmundson, A., Rodgers, G. W., Woodfield, T. B. F., Hooper, G. J., & Chase, J. G. (2012, 29-31 Aug 2012). *Tissue attenuation characteristics of acoustic emission signals for wear and degradation of total hip arthroplasty implants.* Paper presented at the 8th IFAC Symposium on Biological and Medical Systems (BMS12), Budapest, Hungary

Kohn, D. H. (1995). Acoustic emission and nondestructive evaluation of biomaterials and tissues. *Crit Rev Biomed Eng, 23*(3-4), 221-306. doi: 10.1615/critrevbiomedeng.v23.i3-4.20

Kop, A. M., & Swarts, E. (2009). Corrosion of a hip stem with a modular neck taper junction: a retrieval study of 16 cases. *J Arthroplasty, 24*(7), 1019-1023. doi: 10.1016/j.arth.2008.09.009

Kurtz, S., Ong, K., Lau, E., Mowat, F., & Halpern, M. (2007). Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030. *J Bone Joint Surg Am, 89*(4), 780-785. doi: 10.2106/JBJS.F.00222

Kurtz, S. M., Lau, E., Ong, K., Zhao, K., Kelly, M., & Bozic, K. J. (2009). Future young patient demand for primary and revision joint replacement: national projections from 2010 to 2030. *Clinical Orthopaedics and Related Research®, 467*(10), 2606-2612. doi: 10.1007/s11999-009-0834-6

Levy, Y. D., Munir, S., Donohoo, S., & Walter, W. L. (2015). Review on squeaking hips. *World J Orthop, 6*(10), 812-820. doi: 10.5312/wjo.v6.i10.812

Lewis, C. L. (2010). Extra-articular Snapping Hip: A Literature Review. *Sports health, 2*(3), 186-190. doi: 10.1177/1941738109357298

Liu, X. W., Zi, Y., Xiang, L. B., & Wang, Y. (2015). Total hip arthroplasty: areview of advances, advantages and limitations. *Int J Clin Exp Med, 8*(1), 27-36.

Mavrogordato, M., Taylor, M., Taylor, A., & Browne, M. (2011). Real time monitoring of progressive damage during loading of a simplified total hip stem construct using embedded acoustic emission sensors. *Med Eng Phys, 33*(4), 395-406. doi: 10.1016/j.medengphy.2010.10.025

Miller, T. T. (2012). Imaging of hip arthroplasty. *Eur J Radiol, 81*(12), 3802-3812. doi: 10.1016/j.ejrad.2011.03.103

National Services Scotland. (2016). Scottish Arthroplasty Project Biennial Report 2016. Retrieved from: http://www.arthro.scot.nhs.uk/Reports/Main.html

Nevelos, J., Ingham, E., Doyle, C., Streicher, R., Nevelos, A., Walter, W., & Fisher, J. (2000). Microseparation of the centers of alumina-alumina artificial hip joints during simulator testing produces clinically relevant wear rates and patterns. *J Arthroplasty, 15*(6), 793-795. doi: 10.1054/arth.2000.8100

New Zealand Orthopaedic Association. (2017). *The New Zealand Joint Registry Eighteen Year Report January 1999 to December 2016* A. Rothwell, C. Frampton, & T. Hobbs (Eds.), Retrieved from https://nzoa.org.nz/nz-joint-registry

New Zealand Orthopaedics Association. (2015). *The New Zealand Joint Registry Sixteen Year Report January 1999 to December 2014* A. Rothwell, T. Hobbs, & C. Frampton (Eds.), Retrieved from http://nzoa.org.nz/nz-joint-registry

Ng, E. T., & Qi, G. (2001). Material fatigue behavior characterization using the wavelet-based AE technique – a case study of acrylic bone cement. *Engineering Fracture Mechanics, 68*(13), 1477-1492. doi: 10.1016/s0013-7944(01)00040-6

Ohtsu, M., Enoki, M., Mizutani, Y., & Shigeishi, M. (2016). Principles of the Acoustic Emission (AE) Method and Signal Processing. In T. J. Society for Non-Destructive Inspection (Ed.), *Practical Acoustic Emission Testing* (pp. 5-34). Tokyo: Springer Japan.

OpenStax. (2013a). Anatomy of Selected Synovial Joints. In OpenStax Anatomy and Physiology (Ed.), Anatomy & Physiology: OpenStax CNX. Jun 29, 2013. Retrieved from http://cnx.org/contents/4770f844-6eb0-40bf-96c1-888459ce5219@4.

OpenStax. (2013b). Bones of the Lower Limb. In OpenStax Anatomy and Physiology (Ed.), Anatomy & Physiology: OpenStax CNX. Jun 28, 2013. Retrieved from http://cnx.org/contents/738a2420-a409-4190-89d1-2d89d4496a4b@4.

OpenStax. (2013c). The Pelvic Girdle and Pelvis. In OpenStax Anatomy and Physiology (Ed.), Anatomy & Physiology: OpenStax CNX. Jun 5, 2013. Retrieved from http://cnx.org/contents/53a1e895-7e9c-4246-b0e9-77c97a2d4825@3.

OpenStax. (2013d). Synovial Joints. In OpenStax Anatomy and Physiology (Ed.), Anatomy & Physiology: OpenStax CNX. Jun 28, 2013. Retrieved from http://cnx.org/contents/6c5b58ca-6c6d-4ab3-8b41-60c8733a3144@4.

OpenStax. (2013e). Types of Body Movements. In OpenStax Anatomy and Physiology (Ed.), Anatomy & Physiology: OpenStax CNX. Jun 5, 2013. Retrieved from http://cnx.org/contents/a829ec63-2bac-4a7c-aaf4-87a10e71c27f@3.

OpenStax. (2017). Anatomical Terminology. In OpenStax Anatomy and Physiology (Ed.), Anatomy & Physiology: OpenStax CNX. Jul 19, 2017. Retrieved from http://cnx.org/contents/17e4eea8-a005-45af-b835-f756a014cd48@5.

Oppenheim, A. V., & Schafer, R. W. (2004). From frequency to quefrency: A history of the Cepstrum. *IEEE Signal Processing Magazine, 21*(5), 95-+. doi: 10.1109/Msp.2004.1328092

Owen, D., Russell, N., Chia, A., & Thomas, M. (2014). The natural history of ceramic-on-ceramic prosthetic hip squeak and its impact on patients. *European Journal of Orthopaedic Surgery & Traumatology, 24*(1), 57-61. doi: 10.1007/s00590-012-1142-5

Owen, D. H., Russell, N. C., Smith, P. N., & Walter, W. L. (2014). An estimation of the incidence of squeaking and revision surgery for squeaking in ceramic-on-ceramic total hip replacement: a meta-analysis and report from the Australian Orthopaedic

Association National Joint Registry. *Bone Joint J, 96-B*(2), 181-187. doi: 10.1302/0301-620X.96B2.32784

Parvizi, J., Adeli, B., Wong, J. C., Restrepo, C., & Rothman, R. H. (2011). A squeaky reputation: the problem may be design-dependent. *Clinical Orthopaedics and Related Research®, 469*(6), 1598-1605. doi: 10.1007/s11999-011-1777-2

Parvizi, J., Picinic, E., & Sharkey, P. F. (2008). Revision total hip arthroplasty for instability: surgical techniques and principles. *J Bone Joint Surg Am, 90*(5), 1134-1142.

Paul, J. P. (1976). Force actions transmitted by joints in the human body. *Proc R Soc Lond B Biol Sci, 192*(1107), 163-172. doi: 10.1098/rspb.1976.0004

Piriou, P., Ouenzerfi, G., Migaud, H., Renault, E., Massi, F., & Serrault, M. (2016). A numerical model to reproduce squeaking of ceramic-on-ceramic total hip arthroplasty. Influence of design and material. *Orthop Traumatol Surg Res, 102*(4 Suppl), S229-234. doi: 10.1016/j.otsr.2016.03.005

Pivec, R., Johnson, A. J., Mears, S. C., & Mont, M. A. (2012). Hip arthroplasty. *The Lancet, 380*(9855), 1768-1777. doi: 10.1016/s0140-6736(12)60607-2

Protopapas, M. G., & Cymet, T. C. (2002). Joint cracking and popping: understanding noises that accompany articular release. *The Journal of the American Osteopathic Association, 102*(5), 283-287. doi: 10.7556/jaoa.2002.102.5.283

Rabiei, A. (2009). Hip Prosthesis. In R. Narayan (Ed.), *Biomedical Materials* (pp. 349-369): Springer US.

Randall, R. B. (2013). *A history of cepstrum analysis and its application to mechanical problems*. Paper presented at the International Conference at Institute of Technology of Chartres, France.

Rao, A. (1990). Acoustic emission and signal analysis. *Defence Science Journal, 40*(1), 55.

Rashid, M. S., & Pullin, R. (2014). The sound of orthopaedic surgery--the application of acoustic emission technology in orthopaedic surgery: a review. *European Journal of Orthopaedic Surgery & Traumatology, 24*(1), 1-6. doi: 10.1007/s00590-012-1139-0

Restrepo, C., Parvizi, J., Kurtz, S. M., Sharkey, P. F., Hozack, W. J., & Rothman, R. H. (2008). The noisy ceramic hip: is component malpositioning the cause? *J Arthroplasty, 23*(5), 643-649. doi: 10.1016/j.arth.2008.04.001

Restrepo, C., Post, Z. D., Kai, B., & Hozack, W. J. (2010). The effect of stem design on the prevalence of squeaking following ceramic-on-ceramic bearing total hip arthroplasty. *J Bone Joint Surg Am, 92*(3), 550-557. doi: 10.2106/JBJS.H.01326

Rieger, J. S., Jaeger, S., Schuld, C., Kretzer, J. P., & Bitsch, R. G. (2013). A vibrational technique for diagnosing loosened total hip endoprostheses: an experimental sawbone study. *Med Eng Phys, 35*(3), 329-337. doi: 10.1016/j.medengphy.2012.05.007

Rodgers, G. W., Welsh, R., King, L. J., FitzPatrick, A. J., Woodfield, T. B. F., & Hooper, G. J. (2017). Signal processing and event detection of hip implant acoustic emissions. *Control Engineering Practice, 58*, 287-297. doi: 10.1016/j.conengprac.2016.09.013

Rodgers, G. W., Young, J. L., Fields, A. V., Shearer, R. Z., Woodfield, T. B. F., Hooper, G. J., & Chase, J. G. (2014). Acoustic Emission Monitoring of Total Hip Arthroplasty Implants. *IFAC Proceedings Volumes, 47*(3), 4796-4800. doi: 10.3182/20140824-6-za-1003.00928

Roffe, L., FitzPatrick, A. J., Rodgers, G. W., Woodfield, T. B. F., & Hooper, G. J. (2017). Squeaking in ceramic-on-ceramic hips: No evidence of contribution from the trunnion morse taper. *J Orthop Res, 35*(8), 1793-1798. doi: 10.1002/jor.23458

Roguin, A. (2006). Rene Theophile Hyacinthe Laënnec (1781–1826): The Man Behind the Stethoscope. *Clinical Medicine and Research, 4*(3), 230-235.

Ruther, C., Nierath, H., Ewald, H., Cunningham, J. L., Mittelmeier, W., Bader, R., & Kluess, D. (2013). Investigation of an acoustic-mechanical method to detect implant loosening. *Med Eng Phys, 35*(11), 1669-1675. doi: 10.1016/j.medengphy.2013.06.004

Sariali, E., Jin, Z., Stewart, T., & Fisher, J. (2012). Spectral characterization of squeaking in ceramic-on-ceramic total hip arthroplasty: comparison of in vitro and in vivo values. *J Orthop Res, 30*(2), 185-189. doi: 10.1002/jor.21510

Scruby, C. B. (1987). An Introduction to Acoustic-Emission. *Journal of Physics E-Scientific Instruments, 20*(8), 947-953.

Sexton, S. A., Yeung, E., Jackson, M. P., Rajaratnam, S., Martell, J. M., Walter, W. L., . . . Walter, W. K. (2011). The role of patient factors and implant position in squeaking of ceramic-on-ceramic total hip replacements. *J Bone Joint Surg Br, 93*(4), 439-442. doi: 10.1302/0301-620X.93B4.25707

Shark, L.-K., Chen, H., & Goodacre, J. (2010). Discovering Differences in Acoustic Emission Between Healthy and Osteoarthritic Knees Using a Four-Phase Model of Sit-Stand-Sit Movements. *The Open Medical Informatics Journal, 4*, 116-125. doi: 10.2174/1874431101004010116

Sivananthan, S., Goodman, S., & Burke, M. (2014). Failure mechanisms in joint replacement. In P. A. Revell (Ed.), *Joint Replacement Technology* (pp. 370-400): Woodhead Publishing.

Smith, J., Lee, D., Bali, K., Railton, P., Kinniburgh, D., Faris, P., . . . Powell, J. (2014). Does bearing size influence metal ion levels in large-head metal-on-metal total hip arthroplasty? A comparison of three total hip systems. *J Orthop Surg Res, 9*(1), 3. doi: 10.1186/1749-799X-9-3

Staszewski, W., Boller, C., & Tomlinson, G. R. (Eds.). (2004). *Health monitoring of aerospace structures: smart sensor technologies and signal processing*: John Wiley & Sons.

Sugiyama, H., Whiteside, L. A., & Kaiser, A. D. (1989). Examination of rotational fixation of the femoral component in total hip arthroplasty. A mechanical study of micromovement

and acoustic emission. *Clinical Orthopaedics and Related Research®, 249*(249), 122-128.

Taylor, S., Manley, M. T., & Sutton, K. (2007). The role of stripe wear in causing acoustic emissions from alumina ceramic-on-ceramic bearings. *J Arthroplasty, 22*(7 Suppl 3), 47-51. doi: 10.1016/j.arth.2007.05.038

Teague, C. N., Hersek, S., Toreyin, H., Millard-Stafford, M. L., Jones, M. L., Kogler, G. F., . . . Inan, O. T. (2016). Novel Methods for Sensing Acoustical Emissions From the Knee for Wearable Joint Health Assessment. *IEEE Trans Biomed Eng, 63*(8), 1581-1590. doi: 10.1109/TBME.2016.2543226

Tortora, G. J., Derrickson, B. H., Burkett, B., Dye, D., Cooke, J., Diversi, T., . . . Peoples, G. (2016). *Principles of anatomy and physiology* (1st Asia - Pacific ed.). Milton, Qld: John Wiley & Sons.

Ulrich, S. D., Seyler, T. M., Bennett, D., Delanois, R. E., Saleh, K. J., Thongtrangan, I., . . . Mont, M. A. (2008). Total hip arthroplasties: what are the reasons for revision? *Int Orthop, 32*(5), 597-604. doi: 10.1007/s00264-007-0364-3

Walter, W. L., Insley, G. M., Walter, W. K., & Tuke, M. A. (2004). Edge loading in third generation alumina ceramic-on-ceramic bearings: stripe wear. *J Arthroplasty, 19*(4), 402-413. doi: 10.1016/j.arth.2003.09.018

Walter, W. L., Kurtz, S. M., Esposito, C., Hozack, W., Holley, K. G., Garino, J. P., & Tuke, M. A. (2011). Retrieval analysis of squeaking alumina ceramic-on-ceramic bearings. *J Bone Joint Surg Br, 93*(12), 1597-1601. doi: 10.1302/0301-620X.93B12.27529

Walter, W. L., Lusty, P. J., Watson, A., O'Toole, G., Tuke, M. A., Zicat, B., & Walter, W. K. (2006). Stripe Wear and Squeaking in Ceramic Total Hip Bearings. *Seminars in Arthroplasty, 17*(3-4), 190-195. doi: 10.1053/j.sart.2006.09.015

Walter, W. L., O'Toole G, C., Walter, W. K., Ellis, A., & Zicat, B. A. (2007). Squeaking in ceramic-on-ceramic hips: the importance of acetabular component orientation. *J Arthroplasty, 22*(4), 496-503. doi: 10.1016/j.arth.2006.06.018

Walter, W. L., Waters, T. S., Gillies, M., Donohoo, S., Kurtz, S. M., Ranawat, A. S., . . . Tuke, M. A. (2008). Squeaking hips. *J Bone Joint Surg Am, 90 Suppl 4*, 102-111. doi: 10.2106/JBJS.H.00867

Watkins, J., & Mathieson, I. (2009). Chapter 7 - Structure and function of the foot. In J. Watkins & I. Mathieson (Eds.), *The Pocket Podiatry Guide: Functional Anatomy* (pp. 227-301). Edinburgh: Churchill Livingstone.

Weiss, C., Gdaniec, P., Hoffmann, N. P., Hothan, A., Huber, G., & Morlock, M. M. (2010). Squeak in hip endoprosthesis systems: An experimental study and a numerical technique to analyze design variants. *Med Eng Phys, 32*(6), 604-609. doi: 10.1016/j.medengphy.2010.02.006

Weiss, C., Hothan, A., Huber, G., Morlock, M. M., & Hoffmann, N. P. (2012). Friction-induced whirl vibration: root cause of squeaking in total hip arthroplasty. *J Biomech, 45*(2), 297-303. doi: 10.1016/j.jbiomech.2011.10.025

White, R. A., Hughes, M. S., Burd, T., Hamann, J., & Allen, W. C. (2004). A new operative approach in the correction of external coxa saltans: the snapping hip. *Am J Sports Med, 32*(6), 1504-1508. doi: 10.1177/0363546503262189

Whittington, B., Silder, A., Heiderscheit, B., & Thelen, D. G. (2008). The contribution of passive-elastic mechanisms to lower extremity joint kinetics during human walking. *Gait Posture, 27*(4), 628-634. doi: 10.1016/j.gaitpost.2007.08.005

Whittle, M. W. (2014). *Gait analysis: an introduction*: Butterworth-Heinemann.

Wyatt, M., Hooper, G., Frampton, C., & Rothwell, A. (2014). Survival outcomes of cemented compared to uncemented stems in primary total hip replacement. *World J Orthop, 5*(5), 591-596. doi: 10.5312/wjo.v5.i5.591

Zmitrowicz, A. (2006). Wear patterns and laws of wear-a review. *Journal of theoretical and applied mechanics, 44*(2), 219-253.

Appendix A: Conference Proceeding – IFAC 20th World Congress

Cepstrum Analysis for Determining the Fundamental Frequency of Total Hip Replacement Acoustic Emissions

Anthony J. FitzPatrick*, Geoffrey W. Rodgers*, Tim. B.F. Woodfield, and Gary J. Hooper****

**Dept. of Mechanical Engineering, University of Canterbury, Christchurch, New Zealand 8140*
(Tel: 64-3-3642596; e-mail: geoff.rodgers@ canterbury.ac.nz, anthony.fitzpatrick@canterbury.ac.nz).
***Dept. of Orthopaedic Surgery and Musculo-Skeletal Medicine, University of Otago Christchurch, New Zealand*

Abstract: Acoustic emission (AE) monitoring of patients with total hip replacement (THR) implants has potential as a non-invasive and inexpensive diagnostic procedure for identifying and monitoring implant complications such as loosening and implant squeaking. AE signals that are generated from implant interactions can provide insight into implant condition and failure modes through signal characteristics such as frequency content. A substantial database of AEs has been collected from 107 THR patients using an AE monitoring prototype diagnostic device. This manuscript presents the use of the relatively unknown signal processing technique of *cepstrum* analysis as a method for identifying and comparing the frequency content of squeaking THRs from the AE database.

It was observed that squeaking AE signals commonly exhibited frequency spectrums with strong harmonic content and fundamental frequencies that varied over time. Cepstrum analysis has the ability to isolate fundamental frequency content in signals that contain harmonics. The position and transient behaviour of the fundamental frequency content can then be used to correlate and categorise particular AE signals. The cepstrum results presented in this manuscript exhibits an ability to successfully isolate the fundamental frequency content for THR AE squeaks from separate patients. The usefulness of a cepstrum analysis for correlating and contrasting THR AE squeaks is demonstrated. However, the results also demonstrated that when strong harmonics were not present in the AE signals, the cepstrum isolated frequency content of no obvious significance and was even unable to simply identify the main frequency content reliably. Nevertheless, the cepstrum analysis is a promising technique to aid in the analysis of THR AE signals and can provide useful insight into the frequency behaviour of these emissions.

Keywords: Bio-signals analysis and interpretation, Developments in measurement, signal processing, Acoustic emissions, Frequency spectrum, Orthopaedic Implants, Clinical trial

1. INTRODUCTION

An increasing demand for total hip replacement (THR) surgeries has provided a great need for more effective diagnostic procedures for monitoring and identifying wear, loosening, and other failure modes in THR implants. If the condition of THR prostheses can be easily and reliably monitored more frequently, it is hoped that major revision surgeries can be avoided or reduced in complexity through the early detection of complications and wear. As a consequence, the overall cost and resource burden for the health system can be reduced. Research into new diagnostic techniques that are non-invasive, inexpensive, simple to conduct, and free from ionizing radiation currently has great importance in the field of orthopaedics. One such technique of particular interest is the detection and analysis of acoustic emissions.

Acoustic emission (AE) monitoring has been considered an ideal damage monitoring technique due to its sensitivity to deformation processes and its ability to be applied continuously. It can be used to non-invasively and non-destructively monitor both surface and subsurface deformation, fatigue, and fracture of various materials, including soft tissues and metal or ceramic components. The AE technique can also provide insight into the failure mechanisms of materials. Over the past half century the AE monitoring technique has advanced considerably. It has produced some very promising results in orthopaedics by attempting to diagnose various prosthesis complications such as loosening, osteolysis, and infection (Gao *et al.*, 1990; Kohn, 1995; Browne *et al.*, 2005; Glaser *et al.*, 2010; Rieger *et al.*, 2013; Rashid and Pullin, 2014; Piriou *et al.*, 2016).

Recent research has developed an AE monitoring prototype diagnostic device for the assessment of implant wear and stability in THR patients (Rodgers *et al.*, 2014; FitzPatrick *et al.*, 2017; Rodgers *et al.*, 2017). The device utilises ultrasonic transducers attached to the skin surface of the patient and records vibrations that make it to the skin surface as the patient performs prescribed motions. Under local human ethics approval (New Zealand Upper South A regional ethics committee URA/10/11/075), a database of AEs from 107 THR patients has been collected to date, many of whom exhibited implant complications such as audible squeaking or loosening and have since had their implant components replaced. A National Instruments CompacDAQ and NI-9222 analog module recorded each sensor at 100 kHz.

The assessment and analysis of the implant AEs has included the examination of the frequency content present in the AE signals. This manuscript explores some initial findings with regards to the frequency content and characteristics of audible

squeaking and presents the *cepstrum analysis* as a method for identifying and comparing these AEs.

2. CEPSTRUM THEORY

The Cepstrum method is a variation on traditional frequency domain analyses, with some advantages when multiple strong harmonics are present. Bogert et al. (1963) defined the power *cepstrum* of a signal as "the power spectrum of the logarithm of the power spectrum" (Randall, 2013). The concept of a *cepstrum* was developed while trying to identify echoes in seismic signals. Oppenheim and Schafer (2004) state that a signal containing an echo can be represented as

$$x(t) = s(t) + \alpha s(t - \tau), \quad (1)$$

where $s(t)$ is a signal as a function of time t, α is a constant, and τ is the echo delay. The Fourier power spectrum of (1) is

$$|X(f)^2| = |S(f)^2|[1 + \alpha^2 + 2\alpha\cos(2\pi f\tau)], \quad (2)$$

where $|S(f)^2|$ is the power spectrum of $s(t)$ and f is frequency. Hence it can be seen that the spectrum of a signal containing an echo consists of the spectrum of the original signal which modulates a periodic function that is the spectrum contribution of the echo (Oppenheim and Schafer, 2004). Then, by the rules of logarithms, if the logarithm of the power spectrum is taken, the multiplication in (2) becomes the summation of the two components as follows:

$$\log|X(f)^2| = \log|S(f)^2| + \log[1 + \alpha^2 + 2\alpha\cos(2\pi f\tau)]. \quad (3)$$

If (3) is considered as a waveform (i.e. the frequency input f behaves as if it were time and τ as if it were the fundamental frequency of the cosine wave) then this signal simply contains the log spectrum of the original signal and a periodic component which has a 'fundamental frequency' of the echo delay, τ (Oppenheim and Schafer, 2004).

Given that periodic components appear as sharp peaks in the spectrum of a signal, if the Fourier transform of (3) is computed it will contain a peak at the position corresponding to the time that the echo occurred in the original signal, τ (Oppenheim and Schafer, 2004). However, the resulting spectrum, although a Fourier transform of a signal, is not in the frequency domain, as it has a unit of time as the independent variable, but it is also not in the time domain as the signal is different to the original waveform. Thus, Bogert et al. (1963) named this domain the *quefrency* domain and the spectrum of (3) as the *cepstrum* as they found they were "operating on the frequency side in ways customary on the time side and vice versa" (Oppenheim and Schafer, 2004).

While finding the presence of echoes is useful in some applications it is not immediately useful for AE analysis of THR implants. However, cepstrum analysis has another property which is of more immediate use for analysing THR AEs. Cepstrum analysis has the ability of identifying the fundamental frequency of a vibration when multiple harmonics are present. Consider a sinusoid with decayed first and second harmonics of the following form:

$$x(t) = \beta_1 \sin(2\pi f_1 t) + \beta_2 \sin(2\pi f_2 t) + \beta_3 \sin(2\pi f_3 t), \quad (4)$$

where $\beta_1 = 1, \beta_2 = 0.3, \beta_3 = 0.1, f_1 = 2Hz, f_2 = 4Hz$, and $f_3 = 6Hz$. Figure 1a shows the log-magnitude Fourier power spectrum of a Hanning windowed version of the signal defined in (4). As expected, it can be observed from Figure 1a that three frequency peaks are present at the appropriate frequencies of 2, 4, and 6 Hz. By taking the inverse Fourier transform of the log-magnitude signal in Figure 1a the cepstrum in Figure 1b is produced. However, the signal in Figure 1b does not immediately appear useful and so Figure 1c was created by replotting the data of Figure 1b using inverse quefrency values. By using the inverse quefrency values the horizontal axis of Figure 1c has units of Hz and it can be observed that a sharp peak is present at 2 Hz. Hence, the cepstrum has successfully identified the fundamental frequency present in (4) of 2 Hz, despite the strong presence of harmonics.

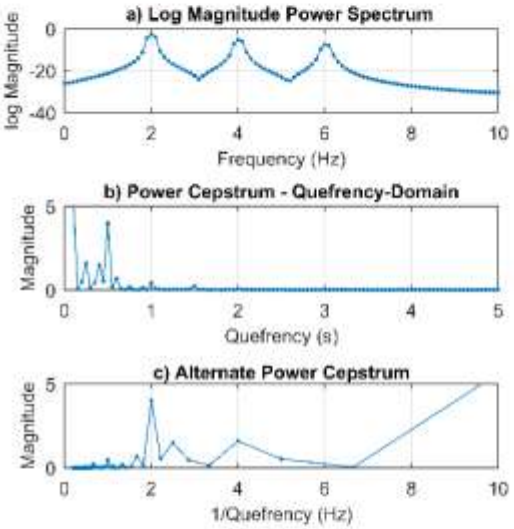

Fig. 1: a) Log-magnitude Fourier power spectrum of a Hanning windowed sinusoid containing two harmonic frequencies, b) cepstrum of the sinusoid, and c) alternative representation of the cepstrum to show units of Hz on the horizontal axis.

The ability to identify the fundamental frequencies of a signal containing harmonics is a useful property of the cepstrum analysis. It had been found from previous work that some AE signals from THR patients, who exhibit audible squeaking, show a strong presence of harmonic frequencies (FitzPatrick *et al.*, 2017). Therefore, this manuscript explores the application of the cepstrum analysis to assist in the comparison and correlation of AE signals of THR patients. Additionally, throughout the duration of audible squeaking, the position of the fundamental frequency tended to show a gradual shift of up to 1 kHz in some cases. As a result, a short-time (ST) cepstrum analysis was applied to better reflect the transient frequency content of the AEs.

3. METHODS

The existing THR AE signals have been sampled at a rate of 100 kHz giving a signal time-step of 0.01ms. To perform the

ST cepstrum analysis MATLAB software was used to apply an algorithm to individual THR AE signals to generate corresponding ST cepstrum signals. The algorithm to compute a ST cepstrum signal was as follows:

1. Divide the signal array into 10ms windows of time and treat each time window as a separate signal between 0 and 10ms (That is, each signal window's time axis array will be from 0 to 10ms in steps of 0.01ms.).
2. Perform a Fast Fourier Transform (FFT) on the first time-domain window signal and compute the modulus of the resulting signal. Note that this new signal array is the frequency-domain response of the time-domain signal and has a corresponding frequency value array from 0 to 100 kHz in steps of 100 Hz ($dF = 100kHz/1000samples = 100Hz$).
3. Compute the natural logarithm of FFT result from Step 2.
4. Perform an inverse FFT of the result from Step 3.
5. Compute the modulus of the result from Step 4. This new signal array is the quefrency-domain response and because a FFT and consequent inverse FFT have been performed, the corresponding quefrency value array is the same as the time-domain time-step values from 0 to 10ms with a quefrency-step of 0.01ms.
6. Repeat Step 2 to 5 for each of the remaining time-domain signal windows.
7. Calculate the array of equivalent frequency values by taking the inverse of the quefrency value array (i.e. $1/quefrency$).

To present the results a 3-dimensional ST frequency-domain plot (known as a spectrogram) and 3-dimensional ST cepstrum plot were created for each AE signal.

The spectrogram was created by plotting each frequency magnitude result array (Step 2 of the algorithm) on the z (out-of-page) axis against the corresponding frequency values on the y (vertical) axis and at the appropriate window position in time on the x (horizontal) axis. The window positions on the x-axis corresponded to the time at the centre of each window from the original time-domain AE signal (for example, the result arrays for the first, second, and third signal windows were placed at 5, 15, and 25ms on the x-axis respectively).

The ST cepstrum was created by plotting each quefrency magnitude result array (Step 5 of the algorithm) on the z-axis against the corresponding equivalent frequency values (Step 7 of the algorithm) on the y-axis and at the appropriate window position in time on the x-axis. The resulting 3D plots are shown here as 2D colormaps.

4. RESULTS

Figure 2a shows the time-domain response of a typical AE signal recording from a THR patient where audible squeaking was present. In general, audible squeaking was observed in the AE signals as regions which exhibited large amplitudes for long relative durations. In Figure 2a substantial audible squeaking occurred between approximately 3.8 and 4.6 seconds, with a much less significant audible squeak occurring between approximately 3.2 and 3.4 seconds. The AE that occurs between 1 and 1.3 seconds of Figure 2a was not an audible squeak but these types of emissions were a frequent occurrence in THR AE signals.

Figure 2b shows the ST frequency-domain response (spectrogram) of the AE signal from Figure 2a with lighter colours representing higher magnitude of frequency content and red dot markers indicating the largest magnitude frequency content. The spectrogram shows the frequency content of the time-domain signal over time and hence any transient frequency behaviour can be identified. For the purpose of this paper the absolute magnitudes of the frequency responses (and quefrency responses) are not important and therefore the absolute magnitude scale (colorbar) has been included only to indicate the relative magnitude of one colour to the next; this analysis is primarily concerned with the locations of peak frequency content and fundamental frequencies which can be found by considering relative magnitudes.

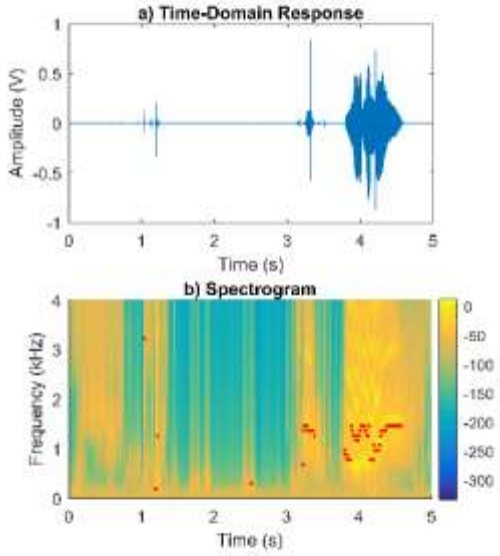

Fig. 2: a) Time-domain response of an AE signal recording from a THR patient showing a large audible squeak at approximately 4 seconds, and b) the corresponding frequency-domain response. Lighter colours representing higher magnitude and peak magnitudes are denoted by red markers.

Figure 3a and 3b show the time-domain response and spectrogram respectively for the same AE signal from Figure 2 but restricted to the time range of 3.7-4.7 seconds that surrounded the large audible squeak. It was observed from Figure 3b that there was a strong presence of harmonics in the signal and the fundamental frequency changes substantially over time (having a range of approximately 1 kHz from approximately 0.7 to 1.7kHz).

The frequency of largest Fourier magnitude (indicated by red dot markers) in Figure 3b was observed to move from the fundamental frequency to the first harmonic for a short time

and then return to the fundamental frequency. This shifting of the largest magnitude frequency between the fundamental and associated harmonics was observed in a number of the THR AEs when audible squeaking was present. This shifting between harmonics creates difficulties in the identification of both the frequency of the fundamental vibration mode and the associated variation in this frequency over time

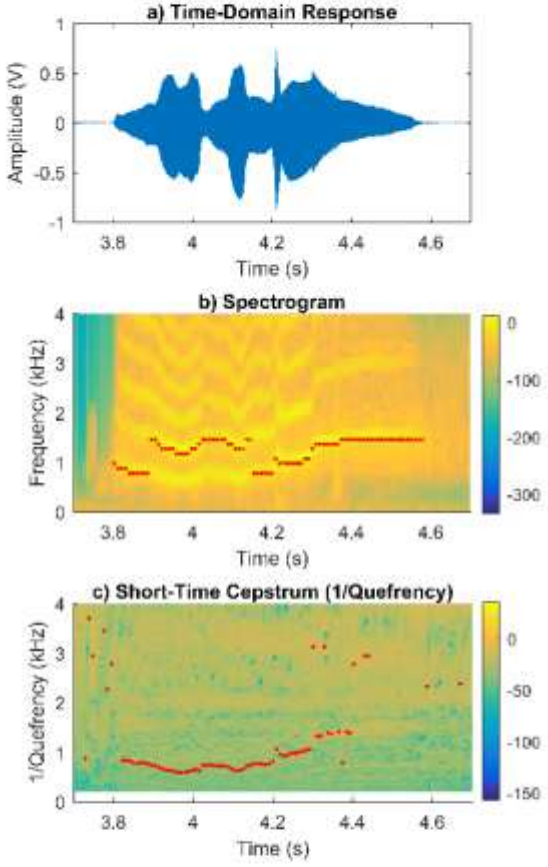

Fig. 3: (a) Time-domain response of an AE audible squeak from a THR patient; (b) corresponding Frequency-domain and (c) quefrency-domain responses showing lighter colours representing higher magnitude and peak magnitudes denoted by red dot markers.

Figure 3c shows the ST cepstrum response of the same AE of Figure 3a and 3b, again with lighter colours representing larger magnitudes. Figure 3c shows that the peak magnitudes, indicated by red dot markers, compared very well with the fundamental frequency peak of the corresponding spectrogram of Figure 3c during the period when strong harmonics were present. During the time windows where the harmonics were less pronounced or non-existent the cepstrum peak magnitudes appeared to demonstrate much higher variability.

Figure 4a shows the time-domain response for an AE audible squeak from a second THR patient and Figure 4b shows the corresponding spectrogram. The spectrogram shows that the majority of the peak frequency content (indicated by the red dot markers) was scattered between 2 to 3 kHz. However, between 2.2 and 2.4 seconds the spectrogram shows the fundamental frequency linearly decreases from approximately 1.2 kHz to approximately 0.2 kHz with the presence of a number of harmonics. Therefore the frequency of largest Fourier magnitude in Figure 4b alone does not necessarily give an accurate representation of the frequency behaviour of the AE, as illustrated by the scattered red dot markers.

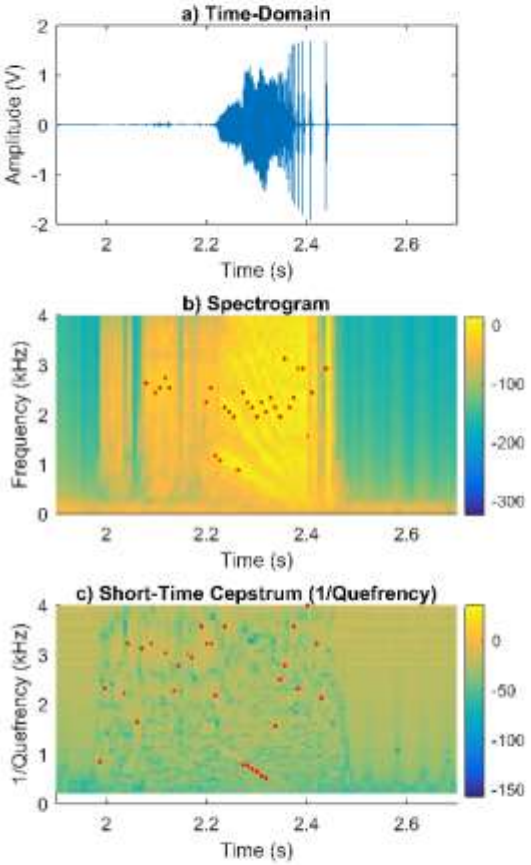

Fig. 4: (a) Time-domain plot of an audible AE squeak from a second THR patient; (b) corresponding frequency-domain and (c) quefrency-domain responses. Lighter colours represent larger magnitude and red dots denote peak magnitudes.

Figure 4c shows the ST cepstrum response of the AE from Figure 4a. It was observed that for regions of the signals where harmonics were not strong or well-defined, the peak cepstrum magnitudes appeared scattered across the quefrency axis. However, in the region between approximately 2.25 and 2.35 seconds the peak cepstrum magnitudes showed a linear decrease from approximately 1 kHz to approximately 0.5 kHz.

Figure 5a shows the time-domain response for an AE audible squeak from a third THR patient and Figure 5b shows the corresponding spectrogram. It was observed from the spectrogram that the largest magnitude frequency between

approximately 0.7 and 0.9 seconds was approximately 3 kHz but showed a small increasing trend with time. The main frequency content from 0.9 seconds dropped to approximately 1.7 kHz and was steady for the remainder of the squeak.

Figure 5c shows the ST cepstrum response of the AE from Figure 5a. It was observed that between approximately 0.8 and 0.9 seconds the peak cepstrum magnitudes were approximately 1.5 kHz and showed a similar shallow increasing trend to that of the peak spectrogram frequencies during the same time period.

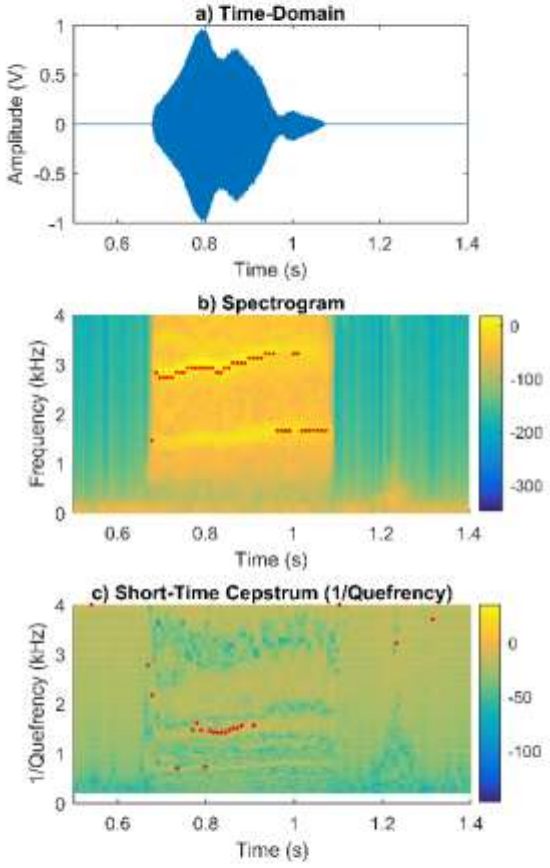

Fig. 5: (a) Time-domain plot of an audible squeak from a third THR patient, (b) corresponding frequency-domain and (c) quefrency-domain plots. Red markers are peak magnitudes.

5. DISCUSSION

The spectrogram results of the three different THR patient's audible squeak signals (Figures 3b, 4b, and 5b) had peak frequency content that varied substantially. Some of the variation in peak frequency content was observed to be due to peak magnitudes switching between the fundamental frequency mode and various harmonics over time. This behaviour can be explained by considering that a larger amount of energy would be required to excite a higher vibration mode of a component and hence if different amounts of vibrational energy are present at different stages of the squeak different vibration modes would be excited more strongly. To support this idea, Figures 5a and 5b show that the period of time after the peak frequency content drops from the 1st harmonic (which was approximately 3 kHz) to the fundamental frequency (which was approximately 1.5 kHz), the time-domain amplitude is significantly reduced and hence less vibrational energy is present.

During the time periods when the peak frequency content of the AE signals was consistently identifying either the fundamental or a given harmonic, some variation in the peak frequency was still observed. This variation was generally limited to a maximum of 200 Hz between time windows. One possible explanation for this additional variation is the influence of implant positioning, muscle forces and associated implant loads during the patient gait cycle. The squeaking signals were recorded while the patient was performing basic orthopaedic test motions such as walking, standing and sitting, and stair ascents. Therefore it is very likely that implant joint angles and loadings changed significantly during the squeak. Literature concerned with squeaking of THRs has found that squeaking frequencies tend to increase with increased implant loads (Hothan *et al.*, 2011). An increase in implant loadings may affect factors such as the stiffness of the entire implant-bone system and consequently modify the natural vibration frequencies of the system.

The results from the ST cepstrums showed that when strong harmonics were present in the AE signals, the peak magnitudes of the cepstrums were closely aligned with the fundamental frequencies seen in the spectrogram results. During these periods where strong harmonics are present, the cepstrum analysis performed well at collapsing the higher harmonics of the signal and identifying the fundamental vibration frequency. The identification of the fundamental vibration frequency allows for more reliable comparisons of squeaking frequencies between separate AE signals. Prior comparisons of AE signals compared the peak magnitude frequency content which exhibited substantial variations due to the harmonics present in the signal. Comparisons between AE signals using the frequencies identified by the cepstrum analysis has the potential to more reliably categorise AE signals based on the fundamental mode and thus assist in identifying signals that characterise particular implant failure modes or conditions.

One of the main limitations of the cepstrum is its inability to isolate frequencies of significance when strong harmonics are not present. This limitation is easily observed in the scattered peak magnitude markers of Figures 3c, 4c, and 5c for periods of time where the corresponding spectrograms show weak harmonic frequencies. It might be expected that the cepstrum would be able to identify the peak magnitude frequency content during time periods without harmonics; however, this was not the case and the associated frequencies of the peak cepstrum magnitudes during these periods had no obvious significance. It is therefore important to carefully consider the frequency content of any signal before applying a cepstrum analysis in order to interpret the results accordingly.

Another limitation of the cepstrum is that while the quefrency resolution is constant, the equivalent frequency resolution is

not constant. The equivalent frequencies of the cepstrum are related to the quefrencies through an inverse relationship (i.e. frequency = 1/quefrency) and hence the frequency resolution of the cepstrum is very high at low frequency and decreases as frequency increases. This trend means that the cepstrum may produce distorted result if the frequencies of interest have a large bandwidth. In the case of the audible squeaks of Figures 3, 4, and 5 a large bandwidth is not a significant issue as the frequencies of interest are below 4 kHz. The coarsest frequency resolution up to 4 kHz is approximately 150 Hz which is sufficient for useful frequency comparisons between squeaks. However, the equivalent frequency resolution must be considered if the frequencies of interest become higher than 4 kHz, or if the length of the cepstrum time window decreases, as this will also affect the size of the quefrency resolution.

A further limitation of the cepstrum is its sensitivity to signal noise. The cepstrum is defined as the "spectrum of the log-magnitude spectrum of a signal" (Randall, 2013). The spectrum of a signal will inevitably contain some level of noise. When the log-magnitude spectrum is computed and then treated as a linear scale in decibels, any noise will be amplified along with the signal. When the inverse Fourier transform is applied to the log-magnitude spectrum the amplified noise will much more significantly affect the result. Higher sampling rates will also have a similar effect as they will include high frequency noise in the signal, which may have otherwise not been recorded. Much of the scatter observed in Figures 3, 4, and 5 can be attributed to time windows without major AE content, resulting in primarily ambient noise with a low signal-to-noise ratio.

6. CONCLUSION

The cepstrum analysis has the ability to isolate fundamental frequencies in signals that contain strong harmonics. The cepstrum analysis was applied to AE data of total hip replacement implants. The cepstrum analysis method shows promise in identifying the primary vibration mode when strong harmonics were present, but resulted in large variation and poor identification in regions where the harmonics were not present. This method has promise, but it is important to carefully consider the frequency content of any signal in order to properly interpret the results of the cepstrum analysis.

7. ACKNOWLEDGEMENTS

This work was supported by the Medical Technologies Centre of Research Excellence (MedTech CoRE), funded by the Tertiary Education Commission of New Zealand. Funding from a HOPE-Selwyn Foundation Scholarship in Ageing Research is also gratefully acknowledged.

REFERENCES

Bogert, B., Healy, M. and Tukey, J. (1963) 'The quefrency analysis of time series for echoes: cepstrum, pseudo-autocovariance, cross-cepstrum and saphe craking', *Proc. Symp. On Time Series Analysis*. John Wiley and Sons, Inc, pp. 209-243.

Browne, M., Roques, A. and Taylor, A. (2005) 'The acoustic emission technique in orthopaedics - a review', *Journal of Strain Analysis for Engineering Design*, 40(1), pp. 59-79.

FitzPatrick, A. J., Rodgers, G. W., Hooper, G. J. and Woodfield, T. B. F. (2017) 'Development and validation of an acoustic emission device to measure wear in total hip replacements in-vitro and in-vivo', *Biomedical Signal Processing and Control*, 33, pp. 281-288.

Gao, X.-J., Murota, K., Tomita, Y., Ono, M., Higo, Y. and Nunomura, S. (1990) 'Evaluation of the Fixation of Artificial Hip Joint by Acoustic Emission', *Japanese Journal of Applied Physics*, 29(S1), pp. 215-215.

Glaser, D., Komistek, R. D., Cates, H. E. and Mahfouz, M. R. (2010) 'A non-invasive acoustic and vibration analysis technique for evaluation of hip joint conditions', *Journal of Biomechanics*, 43(3), pp. 426-432.

Hothan, A., Huber, G., Weiss, C., Hoffmann, N. and Morlock, M. (2011) 'The influence of component design, bearing clearance and axial load on the squeaking characteristics of ceramic hip articulations', *J Biomech*, 44(5), pp. 837-841.

Kohn, D. H. (1995) 'Acoustic Emission and Nondestructive Evaluation of Biomaterials and Tissues', 23(3-4), pp. 221-306.

Oppenheim, A. V. and Schafer, R. W. (2004) 'From frequency to quefrency: a history of the cepstrum', *IEEE Signal Processing Magazine*, 21(5), pp. 95-106.

Piriou, P., Ouenzerfi, G., Migaud, H., Renault, E., Massi, F. and Serrault, M. (2016) 'A numerical model to reproduce squeaking of ceramic-on-ceramic total hip arthroplasty. Influence of design and material', *Orthopaedics & Traumatology: Surgery & Research*, 102(4, Supplement), pp. S229-S234.

Randall, R. B. (2013) 'A history of cepstrum analysis and its application to mechanical problems', *International Conference at Institute of Technology of Chartres, France*. October. pp. 11-16.

Rashid, M. S. and Pullin, R. (2014) 'The sound of orthopaedic surgery—the application of acoustic emission technology in orthopaedic surgery: a review', *European Journal of Orthopaedic Surgery & Traumatology*, 24(1), pp. 1-6.

Rieger, J. S., Jaeger, S., Schuld, C., Kretzer, J. P. and Bitsch, R. G. (2013) 'A vibrational technique for diagnosing loosened total hip endoprostheses: An experimental sawbone study', *Medical Engineering & Physics*, 35(3), pp. 329-337.

Rodgers, G. W., Welsh, R., King, L. J., FitzPatrick, A. J., Woodfield, T. B. F. and Hooper, G. J. (2017) 'Signal processing and event detection of hip implant acoustic emissions', *Control Engineering Practice*, 58, pp. 287-297.

Rodgers, G. W., Young, J. L., Fields, A. V., Shearer, R. Z., Woodfield, T. B. F., Hooper, G. J. and Chase, J. G. (2014) 'Acoustic Emission Monitoring of Total Hip Arthroplasty Implants', *IFAC Proceedings Volumes*, 47(3), pp. 4796-4800.

Appendix B: KS Test Results Tables – Individual Noise and Loosening Revision Reasons

Table B.1: Kolmogorov-Smirnov test results for event duration and RMS amplitude of noise group participants. A significance level of 1% was used.

Pairing	Duration			RMS Amplitude		
	Reject Null Hypothesis	P value	Test Statistic	Reject Null Hypothesis	P value	Test Statistic
Control-P1:	1	<0.001	0.424	0	0.327	0.142
Control-P10:	0	0.120	0.133	0	0.809	0.072
Control-P23:	0	0.057	0.188	0	0.125	0.166
Control-P25:	0	0.200	0.151	0	0.602	0.108
Control-P27:	1	<0.001	0.384	1	<0.001	0.216
Control-P53:	1	<0.001	0.257	1	0.002	0.154
Control-P58:	1	<0.001	0.232	1	0.004	0.111
Control-P6:	1	<0.001	0.215	1	<0.001	0.144
Control-P60:	0	0.029	0.291	1	<0.001	0.438
Noise-P1:	0	0.032	0.213	0	0.045	0.204
Noise-P10:	1	<0.001	0.240	0	0.101	0.135
Noise-P23:	0	0.181	0.152	0	0.992	0.060
Noise-P25:	1	<0.001	0.287	0	0.012	0.223
Noise-P27:	1	<0.001	0.104	1	<0.001	0.087
Noise-P53:	1	<0.001	0.206	1	0.002	0.149
Noise-P58:	0	0.033	0.085	0	0.010	0.096
Noise-P6:	1	<0.001	0.172	1	<0.001	0.139
Noise-P60:	1	0.006	0.338	1	0.008	0.332
P27-P53:	1	<0.001	0.289	1	<0.001	0.224

Table B.2: Kolmogorov-Smirnov test results for event main frequency content and duration between events of noise group participants. A significance level of 1% was used.

Pairing	Main Frequency Content			Duration between Events		
	Reject Null Hypothesis	P value	Test Statistic	Reject Null Hypothesis	P value	Test Statistic
Control-P1:	1	<0.001	0.578	1	0.006	0.289
Control-P10:	1	<0.001	0.323	1	<0.001	0.288
Control-P23:	1	<0.001	0.703	0	0.028	0.240
Control-P25:	1	<0.001	0.615	1	0.003	0.328
Control-P27:	1	<0.001	0.256	1	<0.001	0.465
Control-P53:	1	<0.001	0.563	0	0.251	0.095
Control-P58:	1	<0.001	0.463	1	<0.001	0.231
Control-P6:	1	<0.001	0.529	1	<0.001	0.319
Control-P60:	1	<0.001	0.401	1	<0.001	0.882
Noise-P1:	1	<0.001	0.510	0	0.184	0.182
Noise-P10:	1	<0.001	0.455	1	<0.001	0.502
Noise-P23:	1	<0.001	0.538	1	0.002	0.301
Noise-P25:	1	<0.001	0.685	1	<0.001	0.598
Noise-P27:	1	<0.001	0.131	1	<0.001	0.153
Noise-P53:	1	<0.001	0.345	1	<0.001	0.323
Noise-P58:	1	<0.001	0.251	1	<0.001	0.197
Noise-P6:	1	<0.001	0.495	0	0.534	0.053
Noise-P60:	0	0.023	0.298	1	<0.001	0.893
P27-P53:	1	<0.001	0.393	1	<0.001	0.473

Table B.3: Kolmogorov-Smirnov test results for event rise time and peak amplitude of noise group participants. A significance level of 1% was used.

Pairing	Rise Time			Peak Amplitude		
	Reject Null Hypothesis	P value	Test Statistic	Reject Null Hypothesis	P value	Test Statistic
Control-P1:	1	<0.001	0.415	0	0.548	0.119
Control-P10:	0	0.173	0.124	0	0.222	0.118
Control-P23:	0	0.706	0.099	0	0.029	0.204
Control-P25:	0	0.530	0.114	0	0.428	0.123
Control-P27:	1	<0.001	0.319	1	<0.001	0.281
Control-P53:	1	<0.001	0.199	1	<0.001	0.254
Control-P58:	1	<0.001	0.312	1	<0.001	0.132
Control-P6:	1	<0.001	0.257	1	<0.001	0.143
Control-P60:	0	0.353	0.186	1	<0.001	0.506
Noise-P1:	0	0.047	0.202	0	0.265	0.148
Noise-P10:	1	<0.001	0.330	1	<0.001	0.265
Noise-P23:	1	<0.001	0.280	0	0.387	0.126
Noise-P25:	1	<0.001	0.292	1	<0.001	0.287
Noise-P27:	1	<0.001	0.091	1	<0.001	0.089
Noise-P53:	1	<0.001	0.365	0	0.085	0.100
Noise-P58:	1	<0.001	0.128	1	0.002	0.112
Noise-P6:	1	0.009	0.105	0	0.018	0.098
Noise-P60:	0	0.033	0.286	1	0.003	0.358
P27-P53:	1	<0.001	0.450	1	0.001	0.159

Table B.4: Kolmogorov-Smirnov test results for number of events per recording duration of noise group participants. A significance level of 1% was used.

Pairing	Number of Events per Recording Duration		
	Reject Null Hypothesis	P value	Test Statistic
Control-P1:	0	0.030	0.441
Control-P10:	0	0.268	0.190
Control-P23:	0	0.284	0.205
Control-P25:	0	0.320	0.176
Control-P27:	1	<0.001	0.650
Control-P53:	1	<0.001	0.400
Control-P58:	1	<0.001	0.513
Control-P6:	1	<0.001	0.800
Control-P60:	1	<0.001	0.534
Noise-P1:	0	0.468	0.260
Noise-P10:	1	<0.001	0.404
Noise-P23:	0	0.014	0.331
Noise-P25:	1	<0.001	0.369
Noise-P27:	1	<0.001	0.453
Noise-P53:	1	0.008	0.275
Noise-P58:	1	0.002	0.293
Noise-P6:	1	<0.001	0.565
Noise-P60:	1	<0.001	0.627
P27-P53:	1	<0.001	0.585

Table B.5: Kolmogorov-Smirnov test results for event duration and RMS amplitude of loosening group participants. A significance level of 1% was used.

Pairing	Duration			RMS Amplitude		
	Reject Null Hypothesis	P value	Test Statistic	Reject Null Hypothesis	P value	Test Statistic
Control-P14:	1	<0.001	0.396	0	0.368	0.078
Control-P15:	1	<0.001	0.362	0	0.047	0.111
Control-P16:	1	0.002	0.139	1	0.001	0.142
Control-P17:	0	0.145	0.183	0	0.497	0.132
Control-P18:	1	<0.001	0.348	1	0.007	0.282
Control-P19:	0	0.060	0.369	0	0.884	0.163
Control-P36:	0	0.141	0.158	0	0.291	0.134
Control-P47:	0	0.023	0.188	1	0.006	0.213
Control-P50:	0	0.895	0.050	0	0.742	0.060
Control-P52:	1	<0.001	0.334	0	0.165	0.102
Control-P54:	1	<0.001	0.282	0	0.018	0.141
Control-P55:	1	<0.001	0.337	0	0.480	0.083
Control-P59:	1	<0.001	0.267	0	0.151	0.086
Control-P62:	1	<0.001	0.260	0	0.040	0.113
Control-P65:	1	<0.001	0.210	1	0.002	0.186
Control-P69:	1	0.005	0.112	0	0.021	0.097
Control-P74:	1	<0.001	0.218	0	0.020	0.140
Control-P76:	0	0.027	0.289	1	<0.001	0.431
Control-P79:	1	<0.001	0.346	0	0.062	0.227
Control-P85:	1	<0.001	0.168	0	0.143	0.064
Control-P90:	1	<0.001	0.287	0	0.133	0.126
Loose-P14:	1	<0.001	0.353	0	0.823	0.051
Loose-P15:	1	<0.001	0.329	0	0.186	0.084
Loose-P16:	1	0.005	0.122	1	<0.001	0.155
Loose-P17:	0	0.161	0.177	0	0.652	0.116
Loose-P18:	1	0.003	0.300	1	0.009	0.274
Loose-P19:	0	0.233	0.287	0	0.899	0.159
Loose-P36:	0	0.432	0.118	0	0.305	0.131
Loose-P47:	0	0.256	0.125	0	0.012	0.198
Loose-P50:	0	0.060	0.111	0	0.666	0.061
Loose-P52:	1	<0.001	0.255	0	0.028	0.129
Loose-P54:	1	<0.001	0.245	0	0.016	0.138
Loose-P55:	1	<0.001	0.262	0	0.156	0.109
Loose-P59:	1	<0.001	0.252	0	0.040	0.101
Loose-P62:	1	<0.001	0.207	0	0.087	0.097
Loose-P65:	0	0.018	0.150	1	0.001	0.187
Loose-P69:	1	0.001	0.113	0	0.062	0.079
Loose-P74:	1	<0.001	0.194	0	0.028	0.130
Loose-P76:	0	0.178	0.215	1	<0.001	0.415
Loose-P79:	1	0.005	0.297	0	0.069	0.222
Loose-P85:	1	<0.001	0.108	0	0.061	0.066
Loose-P90:	1	<0.001	0.254	0	0.027	0.155

Table B.6: Kolmogorov-Smirnov test results for event main frequency content and duration between events of loosening group participants. A significance level of 1% was used.

Pairing	Main Frequency Content			Duration between Events		
	Reject Null Hypothesis	P value	Test Statistic	Reject Null Hypothesis	P value	Test Statistic
Control-P14:	1	<0.001	0.394	0	0.023	0.140
Control-P15:	1	<0.001	0.401	0	0.430	0.076
Control-P16:	1	<0.001	0.205	1	<0.001	0.234
Control-P17:	1	<0.001	0.468	0	0.163	0.213
Control-P18:	1	0.009	0.277	0	0.011	0.313
Control-P19:	0	0.011	0.450	1	<0.001	0.677
Control-P36:	1	<0.001	0.391	1	0.004	0.293
Control-P47:	1	<0.001	0.688	1	<0.001	0.312
Control-P50:	1	<0.001	0.333	0	0.019	0.151
Control-P52:	1	<0.001	0.209	0	0.025	0.164
Control-P54:	1	<0.001	0.711	0	0.258	0.112
Control-P55:	1	<0.001	0.674	0	0.030	0.164
Control-P59:	1	<0.001	0.535	0	0.052	0.113
Control-P62:	1	<0.001	0.200	1	<0.001	0.301
Control-P65:	1	<0.001	0.260	1	0.004	0.210
Control-P69:	1	<0.001	0.537	1	<0.001	0.417
Control-P74:	1	<0.001	0.813	0	0.049	0.139
Control-P76:	1	0.009	0.322	0	0.269	0.279
Control-P79:	1	<0.001	0.656	0	0.027	0.347
Control-P85:	1	<0.001	0.570	1	<0.001	0.187
Control-P90:	1	<0.001	0.295	0	0.219	0.126
Loose-P14:	1	<0.001	0.435	1	<0.001	0.207
Loose-P15:	1	<0.001	0.217	0	0.022	0.123
Loose-P16:	1	<0.001	0.170	1	<0.001	0.161
Loose-P17:	1	<0.001	0.567	0	0.067	0.245
Loose-P18:	1	<0.001	0.467	1	0.002	0.354
Loose-P19:	1	0.004	0.486	1	<0.001	0.698
Loose-P36:	1	<0.001	0.586	1	<0.001	0.383
Loose-P47:	1	<0.001	0.527	1	<0.001	0.400
Loose-P50:	1	<0.001	0.505	0	0.598	0.072
Loose-P52:	1	<0.001	0.434	1	<0.001	0.252
Loose-P54:	1	<0.001	0.561	1	0.001	0.206
Loose-P55:	1	<0.001	0.758	0	0.183	0.119
Loose-P59:	1	<0.001	0.304	0	0.144	0.090
Loose-P62:	1	<0.001	0.431	1	<0.001	0.231
Loose-P65:	1	<0.001	0.409	1	<0.001	0.252
Loose-P69:	1	<0.001	0.306	1	<0.001	0.339
Loose-P74:	1	<0.001	0.685	0	0.556	0.077
Loose-P76:	0	0.015	0.306	0	0.060	0.367
Loose-P79:	1	<0.001	0.597	0	0.016	0.366
Loose-P85:	1	<0.001	0.331	1	<0.001	0.131
Loose-P90:	1	<0.001	0.518	0	0.048	0.159

Table B.7: Kolmogorov-Smirnov test results for event rise time and peak amplitude of loosening group participants. A significance level of 1% was used.

Pairing	Rise Time			Peak Amplitude		
	Reject Null Hypothesis	P value	Test Statistic	Reject Null Hypothesis	P value	Test Statistic
Control-P14:	0	0.047	0.115	0	0.054	0.113
Control-P15:	1	<0.001	0.349	0	0.031	0.116
Control-P16:	0	0.092	0.093	1	<0.001	0.220
Control-P17:	0	0.921	0.088	0	0.525	0.130
Control-P18:	0	0.031	0.243	0	0.187	0.183
Control-P19:	0	0.151	0.316	0	0.566	0.219
Control-P36:	0	0.171	0.152	0	0.329	0.130
Control-P47:	1	0.004	0.221	0	0.550	0.100
Control-P50:	1	<0.001	0.265	1	<0.001	0.179
Control-P52:	1	<0.001	0.253	1	0.006	0.157
Control-P54:	1	<0.001	0.300	0	0.285	0.091
Control-P55:	1	0.006	0.169	0	0.079	0.126
Control-P59:	0	0.812	0.048	0	0.920	0.042
Control-P62:	0	0.088	0.101	1	0.005	0.139
Control-P65:	1	<0.001	0.219	0	0.022	0.151
Control-P69:	1	<0.001	0.212	0	0.097	0.079
Control-P74:	1	<0.001	0.307	0	0.019	0.140
Control-P76:	0	0.020	0.298	0	0.185	0.215
Control-P79:	1	0.004	0.302	0	0.028	0.252
Control-P85:	1	<0.001	0.164	0	0.286	0.055
Control-P90:	0	0.087	0.136	0	0.094	0.134
Loose-P14:	0	0.047	0.111	0	0.036	0.115
Loose-P15:	1	<0.001	0.232	0	0.034	0.110
Loose-P16:	1	<0.001	0.144	1	<0.001	0.221
Loose-P17:	0	0.414	0.140	0	0.485	0.132
Loose-P18:	0	0.747	0.113	0	0.197	0.179
Loose-P19:	0	0.590	0.214	0	0.557	0.219
Loose-P36:	0	0.999	0.051	0	0.401	0.121
Loose-P47:	0	0.279	0.122	0	0.686	0.088
Loose-P50:	1	<0.001	0.375	1	<0.001	0.169
Loose-P52:	1	<0.001	0.190	1	0.003	0.158
Loose-P54:	1	<0.001	0.204	0	0.382	0.080
Loose-P55:	1	<0.001	0.202	0	0.084	0.121
Loose-P59:	1	0.006	0.123	0	0.858	0.044
Loose-P62:	1	0.007	0.129	0	0.015	0.121
Loose-P65:	0	0.046	0.134	0	0.027	0.143
Loose-P69:	0	0.023	0.089	0	0.099	0.073
Loose-P74:	1	<0.001	0.190	1	0.007	0.149
Loose-P76:	0	0.237	0.202	0	0.192	0.211
Loose-P79:	0	0.035	0.243	0	0.023	0.255
Loose-P85:	0	0.175	0.055	0	0.239	0.051
Loose-P90:	0	0.930	0.057	0	0.064	0.138

Table B.8: Kolmogorov-Smirnov test results for number of events per recording duration of loosening group participants. A significance level of 1% was used.

Pairing	Number of Events per Recording Duration		
	Reject Null Hypothesis	P value	Test Statistic
Control-P14:	1	<0.001	0.462
Control-P15:	1	<0.001	0.497
Control-P16:	1	<0.001	0.553
Control-P17:	0	0.946	0.125
Control-P18:	0	0.210	0.268
Control-P19:	1	<0.001	0.561
Control-P36:	0	0.216	0.197
Control-P47:	0	0.204	0.218
Control-P50:	1	0.001	0.330
Control-P52:	1	<0.001	0.332
Control-P54:	1	<0.001	0.363
Control-P55:	0	0.157	0.191
Control-P59:	1	<0.001	0.414
Control-P62:	1	<0.001	0.512
Control-P65:	0	0.059	0.231
Control-P69:	1	0.007	0.301
Control-P74:	1	<0.001	0.452
Control-P76:	1	<0.001	0.450
Control-P79:	1	<0.001	0.391
Control-P85:	1	<0.001	0.460
Control-P90:	0	0.428	0.194
Loose-P14:	0	0.018	0.294
Loose-P15:	0	0.018	0.337
Loose-P16:	1	<0.001	0.415
Loose-P17:	0	0.212	0.251
Loose-P18:	0	0.035	0.357
Loose-P19:	1	<0.001	0.645
Loose-P36:	1	0.002	0.344
Loose-P47:	0	0.034	0.288
Loose-P50:	0	0.191	0.184
Loose-P52:	0	0.050	0.207
Loose-P54:	0	0.029	0.213
Loose-P55:	0	0.650	0.123
Loose-P59:	0	0.031	0.249
Loose-P62:	1	0.002	0.381
Loose-P65:	0	0.211	0.182
Loose-P69:	0	0.043	0.245
Loose-P74:	0	0.038	0.279
Loose-P76:	1	<0.001	0.533
Loose-P79:	1	<0.001	0.516
Loose-P85:	1	<0.001	0.338
Loose-P90:	0	0.039	0.309

Appendix C: Data Collection Instructions

Data Collection Instructions (Software Focussed)

Anthony FitzPatrick

July 2018

1 Introduction

This document is intended to aid in the data collection process for the Hip Assessment and Revision Prediction (HARP) project. The steps outlined here are focussed primarily at using the various different software programs used for the acoustic emission (AE), inertial measurement unit (IMU), and 3D motion capture (MoCap) data collection. The penultimate section of the document is concerned with the initial post-processing of the data using MATLAB to organise and time synchronise the AE and IMU data sets. The final section of the document deals with the process of time synchronising the MoCap data set with the IMU data set.

2 Acoustic Emission

a) Ensure that the four AE sensors are connected to the the NI 9222 module and that the module is attached to the NI cDAQ-9174 (CompacDAQ). Connect the CompacDAQ via USB to the laptop/computer and power on.

b) Start LabVIEW and open the file `AE Sensor DAQ system timestamps.vi`.

c) Click run on the program and choose a save location and filename from the pop-up window. (NOTE: The filename should be either `PatientXXX_1`, `NormalXXX_1`, or `BenchXXX_1` for patient, control, or bench testing respectively where `XXX` is replaced by the appropriate digits for the current set of testing. The initial filename should end with the digit `1`, as shown, so

that when new data is recorded the filename will be incremented by one.)

d) The sensor voltages should now be displayed on the LabVIEW front panel both as digits and on the time-voltage chart. Check that all the sensors respond to input by lightly tapping the sensor housings and identifying the corresponding voltage spikes on the LabVIEW front panel.

e) Attach the AE sensor pad to the participant at the appropriate location.

f) Record AE data by clicking the `Record` button on the LabVIEW front panel. To stop recording, press the `Record` button again. (NOTE: The AE data is sampled at 100 kHz and hence generates a large amount of data. It is not recommended to leave the AE data recording unnecessarily. Each time a new recording is taken a new save file will be generated.)

g) When recording for a participant is complete the LabVIEW program can be stopped. Alternatively, if a new participant is to be tested, the base filename can be changed from the front panel and recording can be continued.

3 Inertial Measurement Units)

3.1 IMU Connection & Python GUI Controls

a) Plug in the four *SparkFun 9DoF Razor IMU M0* IMUs to the laptop USB ports.

b) Run the .bat file with the filename `Run Razor_9DOF_M0_vX2_2.bat`. Running this file should run the IMU data collection python program (`Razor_9DOF_M0_vX2_2.py`) using the windows command processor. It should also be possible to run the python program directly by clicking the `Razor_9DOF_M0_vX2_2.py` file.

c) From the pop-up window (Figure 1), select a save location and enter a base filename for any IMU data. For each IMU, a .csv file will be created with the IMU name appended to the base filename (e.g. `Test.csv` becomes `Test_IMU0.csv`, `Test_IMU1.csv`, etc. See Figure 5).

If you wish to change the base filename you can do so by clicking the *Choose Base Filename*

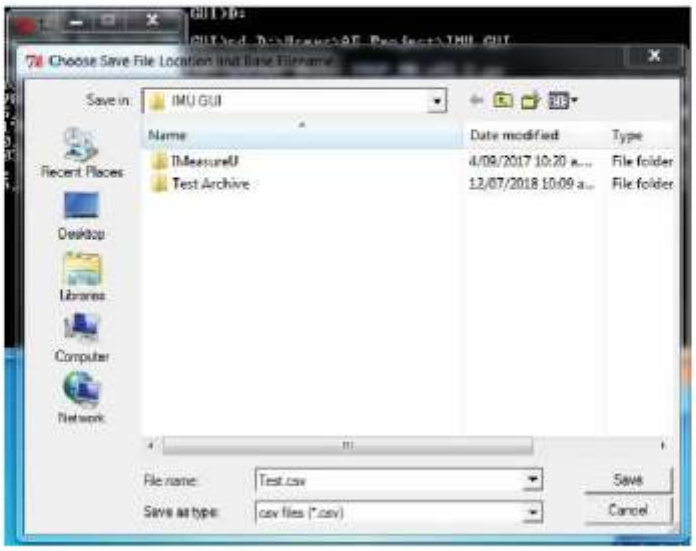

Figure 1: Screenshot of Python program pop-up window asking for a save location and base filename.

button in the main Python GUI window. The save location and base filename pop-up window will appear again upon clicking the *Choose Base Filename* button. Note that the base filename will be updated immediately. However, before save files using the new base filename are created, logging of all IMUs must be stopped momentarily. Save files using the new base filename are then created once logging on any IMU recommences.

d) After entering a base filename the main Python GUI window should appear as in Figure 2 if four IMUs are connected.

Figure 2: Screenshot of Python program main GUI window if four IMUs are connected.

If four IMUs do not appear in the GUI then exit the GUI and check the Windows Device Manager to ensure the four IMUs are connected as four separate COM ports (Figure 3). If so, then restart from **b)**.

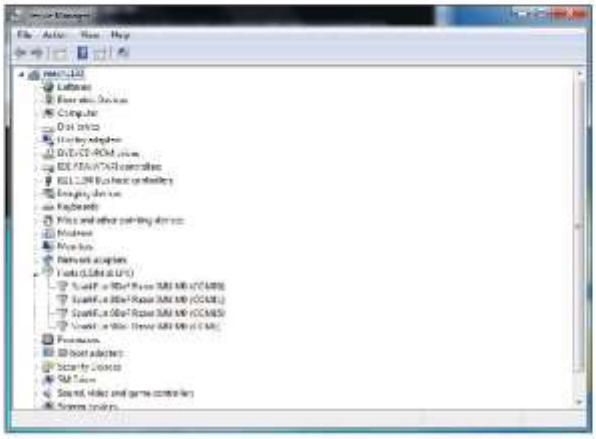

Figure 3: Screenshot of Windows Device Manager showing four IMUs connected to COM ports.

e) To begin logging with an IMU, click the corresponding *Log* button. Once clicked, the button text should change to *Stop* (See Figure 4) and data will be recorded to the corresponding file. If a save file already exists, any new IMU data will be appended to the end of any existing data in the save file.

Figure 4: Screenshot of Python GUI showing three IMUs logging. Note IMU 0 is not logging as its Log button is still displaying Log.

f) To stop logging with an IMU, click the corresponding *Stop* button. Once all IMUs have been stopped from logging the save files will be closed and can therefore be accessed by other programs.

If any IMUs have been logging, a save file for every IMU is created even if the other IMUs have not been logging (See Figure 5). A save file from an IMU that has not been logging will contain only column headers.

Figure 5: Screenshot of IMU save files showing data has been saved for IMUs 1-3. IMU 0 was not commanded to log data and hence no data has been saved (indicated by a 1 KB file size). Therefore, the save file for IMU 0 will only contain the column headers.

3.2 IMU Calibration

It is recommended that calibration data be collected for each IMU at the beginning of every test session or save file. Calibration of the IMUs is done in three stages: determining gyroscope offsets, determining accelerometer offsets and scale factors, and finally determining magnetometer offsets and scale factors. The gyroscope offsets are found by recording data while the IMU is stationary. The accelerometer offsets and scale factors are found by recording stationary data of the IMU with each of its accelerometer axes (both positive and negative) aligned with acceleration due to gravity. The accelerometer calibration assumes the calibration data to have a maximum acceleration of 1 g for each of the six orientations (+X, -X, +Y, -Y, +Z, -Z). The magnetometer offsets and scale factors are found by recording data from the IMU while it is rotated through a number of different orientations in free space. A sphere-fit of the magnetometer calibration data is used to find the appropriate offsets and scale factors. It should be noted that by using a sphere-fit it is assumed that any hard- and soft-iron distortions that may be present have an insignificant effect on the magnetometer. Some initial tests were performed to determine if the aforementioned assumption was reasonable and it was found that the test data produced nominally spherical results.

To collect data for IMU calibration the following steps should be taken:

a) Ensure that the name of each IMU is known. The enclosure of each IMU has been labelled either 0, 1, 2, or 3 (Note that each IMU has been hard-coded with an IMU name, see `Razor_Firmware_V3.ino` and its corresponding `config.h` file, and therefore if any enclosures

have been switched the enclosure labels may not match the IMU names).

b) Place one IMU on a flat surface with the positive Z acceleration axis directed upwards (as shown in Figure 6).

Figure 6: An IMU in the initial calibration position with its positive Z acceleration axis directed upwards and perpendicular to the table surface.

c) With the IMU held stationary, log at least two seconds of data by pressing the appropriate `Log` button on the GUI. Stop logging IMU data by pressing the corresponding `Stop` button on the GUI. (NOTE: Be very careful not to move the IMU before logging has stopped. This part of the calibration data is used to calibrate both the accelerometer and the gyroscope and hence any movement of the IMU will negatively affect both calibrations.)

d) Invert the IMU so that the Z acceleration axis is directed in the opposite direction to that of step **b)** (i.e. orientate the IMU so that the enclosure lid is face-down on the table surface). Repeat step **c)** with the IMU in this new position. (NOTE: Be sure to wait at least two seconds between stopping calibration data collection for one orientation and starting calibration data collection for another orientation. The post-processing calibration algorithm relies on a time gap being present between calibration orientation data recordings in order correctly calibrate the IMUs.)

e) Repeat steps **b)**, **c)**, and **d)** twice more for the X and Y axes each in turn. Figure 7 shows

images of how the table surface and edge can be used to aid the collection of acceleration calibration data for the X and Y axes.

Figure 7: Images indicating various IMU calibration positions in relation to a table surface and edge.

f) At this point, six sets of calibration data should have been collected and saved in a single csv file for the corresponding IMU. The magnetometer calibration data can now be collected. Click the `Log` button for the corresponding IMU and begin to rotate the IMU around in free space near the area where testing will occur (keeping the IMU away from any metal objects or objects which may be generating their own magnetic fields). Try to move the IMU through as many orientations as possible while recording so that enough data is collected for a good sphere-fit (NOTE: The post-processing calibration algorithm relies on gyroscope readings above a threshold to indicate that the IMU is moving and hence when the magnetometer calibration began. Therefore, make sure not to rotate the IMU too slowly during magnetometer calibration data collection.). Once the IMU has been rotated through a large range of orientations, click the corresponding `Stop` button to stop logging calibration data.

g) The calibration data collection procedure should now be complete. Actual test data can now be collected from this point onwards by clicking the corresponding `Log` button. It is recommended that the test data be appended to the same csv file as the calibration data (by keeping the base filename the same throughout calibration and testing).

3.3 IMU Test Data Collection

a) Attach three of the IMUs to the participant in the desired locations. The fourth IMU should be kept separate as it will be used, if motion capture data is also being collected, to collect time synchronisation data.

b) Once the IMUs are attached, begin logging IMU data for each IMU by clicking the appropriate `Log` buttons. Logging can be stopped at the end of testing by clicking the appropriate `Stop` buttons.

c) Upon the completion of testing, stop all IMUs from logging before finally exiting the GUI.

4 BTS Bioengineering SMART Motion Capture System

4.1 Camera Set-up, Calibration, & Data Capture (using SMART Capture)

The procedures for the set-up and calibration of the BTS MoCap system are outlined in the *SMART-D User Manual*. Additionally, the document entitled *MoCap Data Collection* contains more specific details regarding the MoCap data collection for this project. However, some important things to note for the MoCap concerning this project are as follows:

- For ease when synchronising the data it is desirable that both the MoCap data is collected at a sampling rate of 100 Hz. Before commencing the data collection, check that the cameras and also the force platform are sampling at 100 Hz.
- The BTS system zeroes the force platform when a recording is started. Therefore, ensure that the participant is not standing on the force platform when a recording is started. (NOTE: This means that a static recording will not give the participant's mass unless they began the recording off the force platform and then move onto it.)
- When recording walking motions it is advisable that the participant starts a few paces away from the force plate so that when they reach the force platform they are more likely to be walking with their natural gait. This also means the participant should not stop walking until they are at least a full pace past the force platform.

- Each recording will need to be saved as a separate tdf file (immediately after recording) so ensure that each file is saved with an easily identifiable filename before commencing a new recording.
- At the time of writing, the best way to synchronise simultaneously collected MoCap and AE data was to include, at the start of each MoCap recording, three distinct impacts on the force plate that were also detected by one of the IMUs. The IMU acceleration data could then be synchronised with the force platform impacts during post processing. Hence, if no impacts were recorded for each MoCap recording then subsequent synchronisation of the datasets was not possible.

4.2 Marker Labelling

a) Start the *SMART TDF Inspector* and then open the tdf file which is to be processed. Once opened, right click on the *block* labelled `Force Platform Data Sequence` and select *Export to Gr* (See Figure 8) and export the file. (NOTE: This .gr file will be used at a later stage to re-introduce the force platform data which gets lost during the labelling and gap filling of the markers.)

b) Repeat Step **a**) but instead use *Export to Emt*. The *Emt* file will be used during post-processing by MATLAB.

c) Start the *SMART Tracker* program and open the same tdf file from step **a**). Move the frame slider to a frame where all markers are visible and distinguishable (see Figure 9). A good frame to use is usually mid-stride and where one foot is planted on the force plate.

d) Open (or create) the model shown in Figure 10 (filename: `ajf130_Main_Force.xmf`) which has the 29 markers, plus the force platform, used during the MoCap. (NOTE: The marker labels should be as indicated in Figure 10 to make later stages of processing more seamless.)

e) In the viewer window, left click on each marker (and the force vector from the force plate) and label appropriately. Once all markers are labelled (for the current frame) they should appear blue in the model editor window.

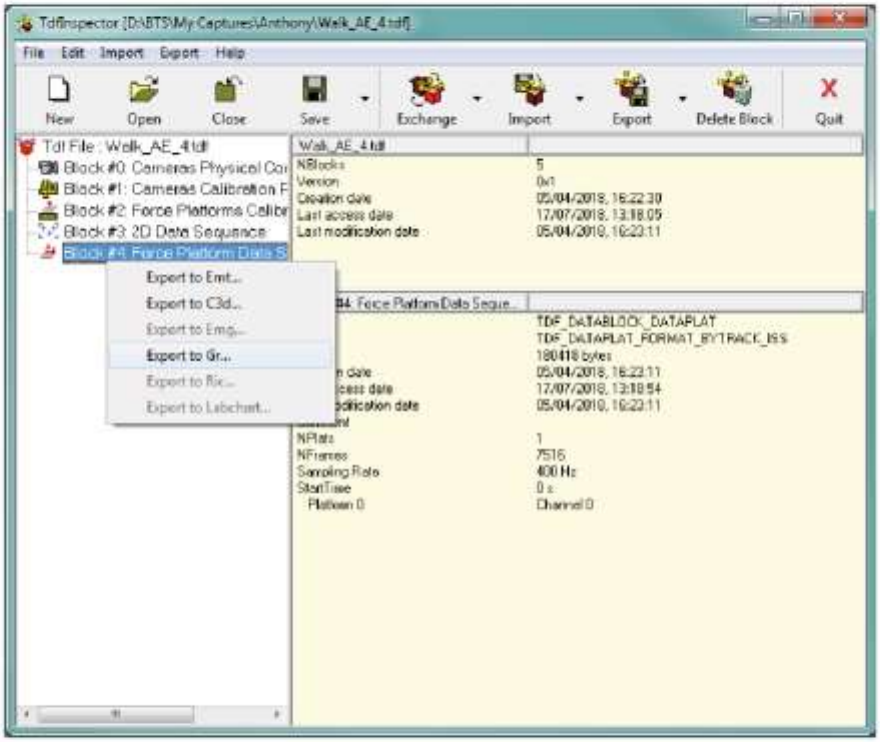

Figure 8: Screenshot showing exporting of the force plate data as a .gr file format.

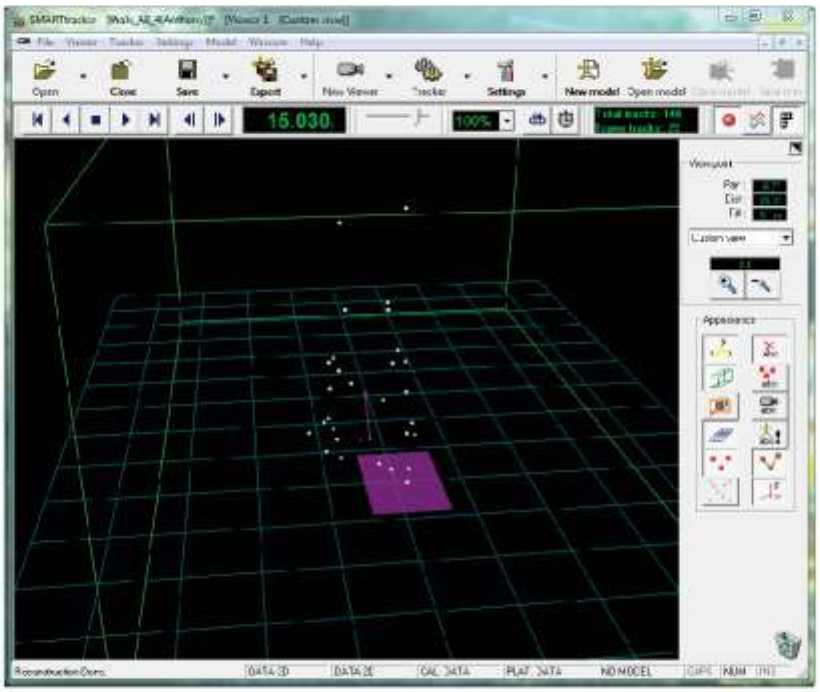

Figure 9: Screenshot showing `SMART tracker` file at a frame where all markers are visible and distinguishable.

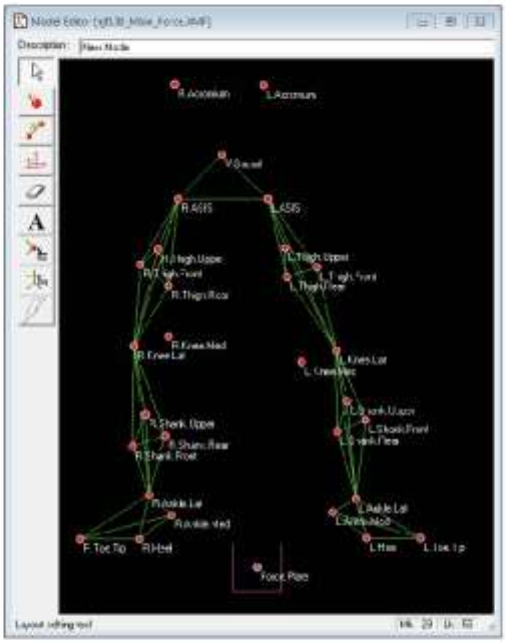

Figure 10: Screenshot of `SMART Tracker` model with 29 markers, plus force platform, used for the MoCap analysis.

At this point it is important to note that the software will have already created what it calls *marker tracks* for each separate marker that it believes was present in the MoCap data. Therefore, when a marker is *labelled* as a particular model marker it is actually the *track* that the marker belongs to that is *labelled*. What this means is that markers can be *labelled* in a single frame and because they are part of a *track*, they will also be *labelled* in any other frame that the particular *track* is in. This should be obvious by moving the frame slider through different frames and observing the labelled and unlabelled markers in the viewer window after step **e**) has been performed.

In a perfect MoCap recording there would be the same number of marker tracks as there were markers used during the recording. However, because the MoCap reconstruction is not perfect, it is highly likely that more marker tracks are present in the *SMART Tracker* program than there were markers used during recording. This generally occurs when a marker is obscured from the cameras for a number of frames and when it reappears the software interprets it as a completely new marker with its own track. Hence, it is necessary to manually check for unlabelled tracks and relabel them appropriately.

f) In the viewer window, move the frame slider and approximately determine the beginning and end times for the *frames of interest*. The *frames of interest* are the frames that include at least one full gait cycle for the leg that traverses the force platform.

g) Open the *graph viewer* by clicking on the *Hide/Show graph viewer* button (See Figure 11) which is found in the top right corner of the viewer window.

Figure 11: SMART Tracker `Hide/Show graph viewer` button.

h) Click on the `Select Tracks` button (See Figure 12) to open the *select tracks* pop-up window. The window should contain all the tracks that have been labelled but will also likely contain a number of unlabelled tracks indicated by square bracketed numbers (See Figure 13). Many of these unlabelled tracks will, in reality, be the tracks of already labelled tracks and hence the unlabelled tracks will need to be merged with the appropriate labelled tracks.

Figure 12: SMART Tracker `Select Tracks` button.

Figure 13: Screenshot of the *select tracks* window showing labelled and unlabelled tracks.

i) To find unlabelled tracks that need to be merged with labelled tracks start by selecting a *like-group* of labelled tracks (i.e. select the three right thigh marker tracks). The *graph viewer* window should now show the positions over time (tracks) of the *like-group*, see the example in Figure 14. The example of Figure 14 shows that data for one of the marker tracks (right thigh rear marker) in the *like-group* is missing for the first part of the *frames of interest*.

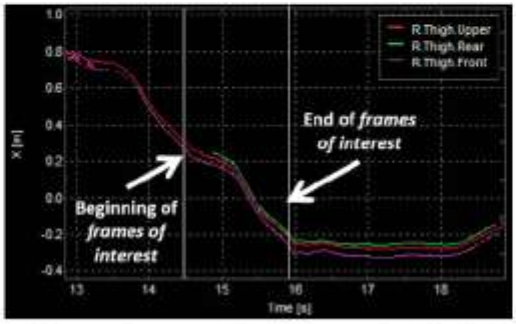

Figure 14: Screenshot of the *graph viewer* window showing the right thigh tracks with *frames of interest* indicated by annotations.

j) If data is missing in the *frames of interest* for any of the marker tracks selected in step **i)**, the unlabelled tracks may contain some of the missing data. To determine if this is the case, move the frame slider so that the vertical white line (which indicates the time of the current frame) of the *graph viewer* is in a region where data is missing.

k) Check the 3D viewer to see if a marker exists that corresponds to the position of the marker that contains missing data (See Figure 15a for an example). If an unlabelled marker is present in the correct position, right click on the unlabelled marker and note the track name assigned to it from the pop-up menu, the track name should be a square bracketed number (See Figure 15b for an example).

l) Reopen the *select tracks* window, Step **h)**, and select the track identified in Step **k)**. The *graph viewer* should now show the positions over time of the unlabelled track. Check that the position values of the unlabelled track are consistent with the trends of the corresponding labelled track. If so, then left click on the unlabelled marker in the 3D viewer and select the

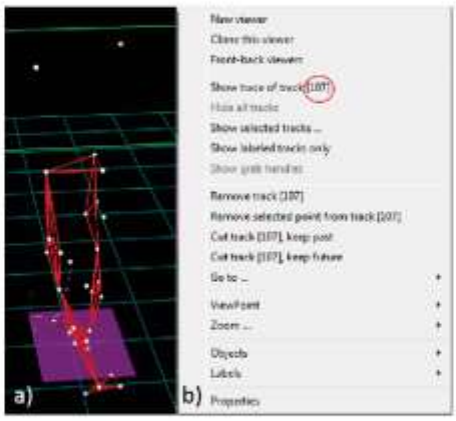

Figure 15: Screenshot of the 3D viewer showing a) an unlabelled marker in the position of the right thigh rear marker, and b) the menu that appears after right clicking on the unlabelled marker showing the track name given to the unlabelled marker (circled), "[107]" in this case.

correct marker name from the list. The unlabelled track should now merge with the correctly labelled track.

m) Repeat from Step **j**) until all unlabelled tracks corresponding to the markers of the *like-group* are appropriately merged. This process only needs to be performed for the unlabelled tracks that are within the *frames of interest* and within approximately 0.5 seconds either side of the *frames of interest*. (NOTE: In order to perform interpolation of tracks (Section 4.4) some data either side of the *frames of interest* is needed.)

n) Repeat from Step **i**), changing the *like-group* each iteration, until the process for all *like-groups* has been completed.

An alternative way to identify if the unlabelled tracks contain useful missing data is to, in the *select tracks* window, select only one of the unlabelled tracks at a time and check if it falls within (or near) the *frames of interest*. If an unlabelled track does fall in (or near) the *frames of interest*, determine its position in the 3D viewer and then select the labelled track, in the *select tracks* window, that the unlabelled track is likely to correspond to. Check the trends of both tracks and, if they are consistent, label the track by left clicking the marker

and selecting the correct marker name.

o) Save the file using *Save as* and ensure that only the labelled tracks are selected to be saved. Ensure that the *selected range* includes all frames, this is import when the force platform data is reintroduced. The options specifically concerning the force track do not matter as the force platform data is taken care of separately. Figure 16 shows the *output data options* pop-up window with the settings to be used when saving. (NOTE: It would be wise not to save over the original file in case something has been done incorrectly.)

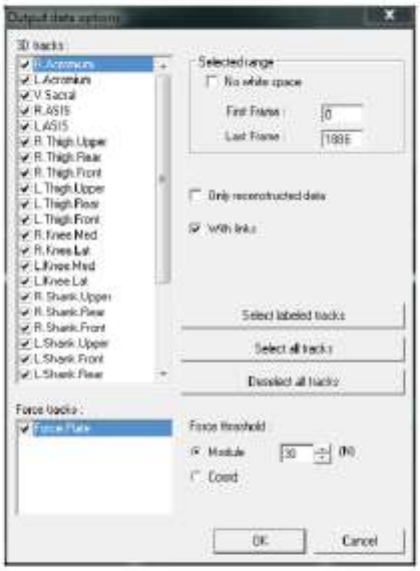

Figure 16: Output options when saving the labelled markers file.

4.3 Remove Erroneous Markers

If any of the labelled tracks contain erroneous markers/frames in the *frames of interest*, such as the one seen in Figure 17, the erroneous markers/frames can be removed by the process outlined in this section. (NOTE: The error shown in Figure 17 was easily identified by observing the change in position of the left knee medial marker in the 3D viewer while moving from the frame identified

in the figure to the next frame. Also, in the erroneous frame, the 3D viewer showed the position of the marker relative to the other markers of the body to be unrealistic.)

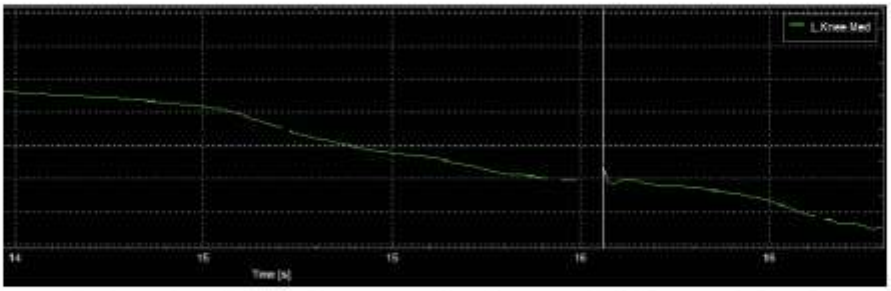

Figure 17: Screenshot of the *graph viewer* window showing the left knee medial track with an erroneous marker/frame at the time indicated by the vertical line.

a) Make sure the file that was saved in Step 4.2**o**) is open in *SMART Tracker* and the appropriate model is also open, Step 4.2**d**).

b) With the erroneous labelled track selected in the *select tracks* window, move the frame slider to the last frame **before** the erroneous frame where the marker is present.

c) In the 3D viewer, right click on the marker that corresponds to the selected track and select the *Cut track [track name], keep past* option from the pop-up menu (Figure 15b). In the *graph viewer* window, the labelled track should have disappeared for all times after the current frame.

d) Reopen the *select tracks* window and select the new unlabelled track (which should be named [0]).

e) Move the frame slider to the first frame **after** the erroneous frame where the marker is present (See Figure 18).

f) In the 3D viewer, right click on the marker that corresponds to the selected track and select the *Cut track [track name], keep future* option from the pop-up menu (Figure 15b). In the *graph viewer* window, the erroneous frame should have disappeared.

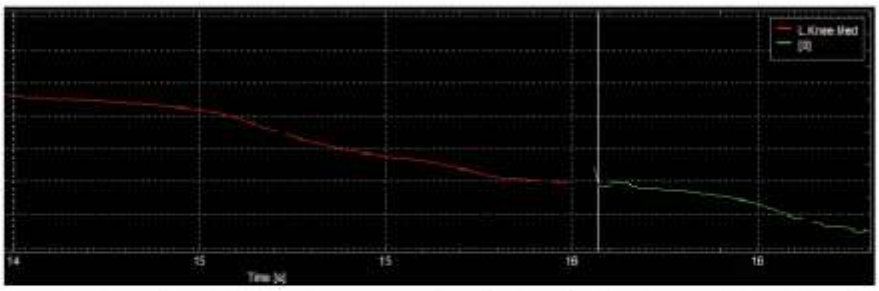

Figure 18: Screenshot of the *graph viewer* window showing the left knee medial track with the current frame (indicated by the vertical line) at the first frame after the erroneous frame. Note the track from the erroneous frame onwards is unlabelled (or rather labelled [0]).

g) In the 3D viewer, left click the marker that corresponds to the unlabelled selected track and select the appropriate marker name. The erroneous frame should now be removed (See Figure 19).

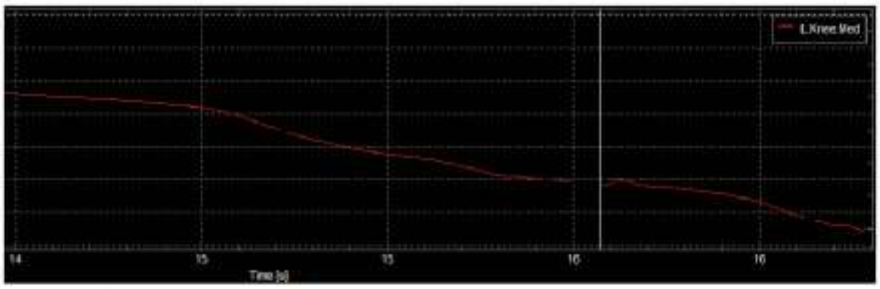

Figure 19: Screenshot of the *graph viewer* window showing the left knee medial track with the erroneous frame removed.

h) Repeat process from Step **b**) for each labelled track that contains erroneous frames.

i) Once all erroneous frames have been removed the file can be saved using the same process as Step 4.2**o**).

Additionally, the file must be exported in C3d format to allow the *SMART Analyzer* to be used to perform marker track interpolation. Select C3d format from the export menu. In the *export data options* pop-up window, ensure all the 3D tracks are selected and that all frames

are included in the selected range.

4.4 Marker Track Interpolation

a) Start the *SMART Analyzer* program, select *Import C3D file* from the file menu, and choose the C3d file exported in Step 4.3**i**). Figure 20 shows the *SMART Analyzer* after the import of the C3d file.

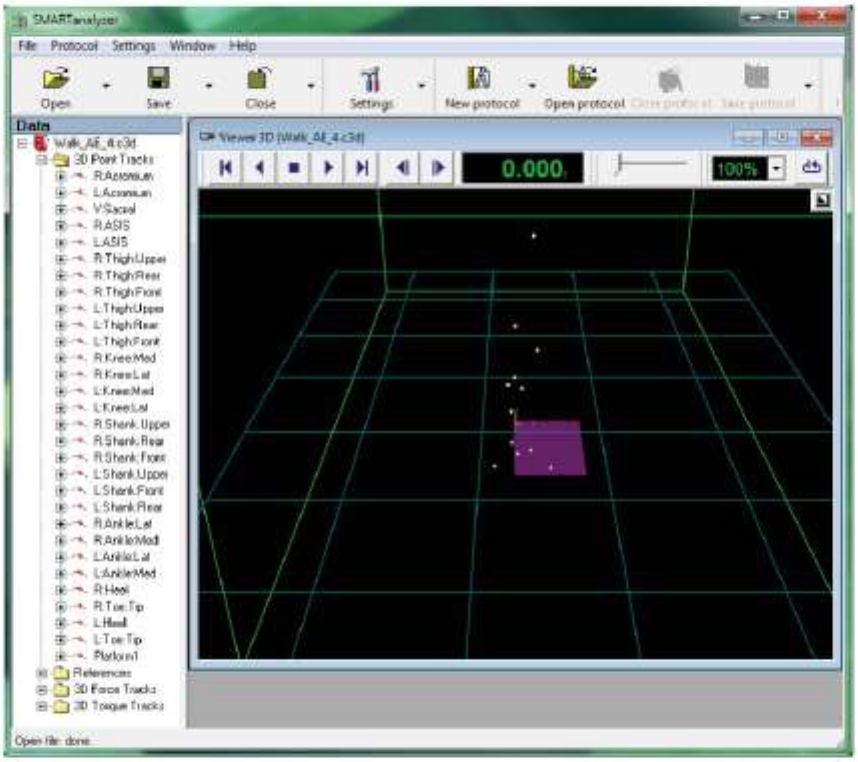

Figure 20: Screenshot of the *SMART Analyzer* after importing the C3d file.

b) Click *Open protocol* and load the protocol with the filename `HARPinterpolation.epx`. A pop-up window named *Sequence association* should appear (See Figure 21). In the *Sequence association* window, click on the blue arrow to apply the interpolation protocol to the C3d

file and click then OK.

Figure 21: Pop-up window after opening a protocol in *SMART Analyzer*.

c) The protocol panel should appear and it should contain 29 spline track interpolation blocks, each block having one of the 29 marker tracks as input and a corresponding interpolated marker track output. See Figure 22 for a sample of what the protocol panel should look like. The names of the interpolated track outputs are the same as their corresponding inputs except with *:Int* appended to the end. Hence, the data panel on the left-hand side of the *SMART Analyzer* window should now contain the interpolated tracks under the *3D Point Tracks* folder (in addition to the original tracks).

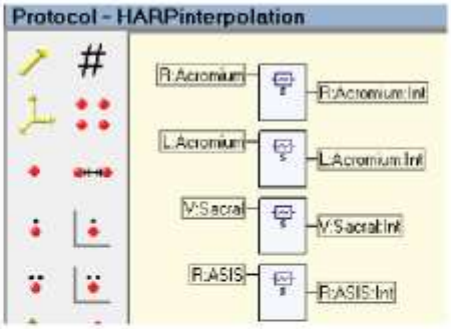

Figure 22: Sample of the protocol panel showing the spline interpolation blocks with marker tracks as input and corresponding interpolated track outputs.

d) Save the file by right clicking on the C3d file in the data panel on the left-hand side of the

SMART Analyzer window and choosing *Save to Tdf* from the menu. In the *Select data* pop-up window, expand the *3D Point Tracks* folder and deselect all the original marker track labels and the *3D Force Tracks* (i.e. only the interpolated point tracks should be selected). Click OK to save the Tdf file and close the *SMART Analyzer*.

e) Start the *SMART Tracker* program and open both the file saved in Step **d**) and the marker model of Step 4.2**d**).

f) Move the frame slider to a time inside the *frames of interest* and then relabel all the markers by left clicking on each marker and selecting the appropriate label from the list.

g) Finally, save the file by clicking save and then export as a C3d file as in Step 4.3**i**) (use a different filename for the C3d file than the one used previously).

4.5 Creating Final TRC File

a) Start the *Mokka* program and import the C3d file of Step 4.4**g**) (Select *File* → *Import* → *C3D file*). Then import the force platform gr file created during Step 4.2**a**) (Select *File* → *Import* → *BTS Bioengineering* → *GR* file*). (NOTE: The visualisation of the position markers of the force platform may need to be turned off by deselecting the position check-box under the *Force platform reactions* collapsible list in the *Acquisition explorer* at the right-hand side of the main window, See Figure 23.)

b) In the *Acquisition explorer*, load the visual configuration file (`Visual Configuration.vst`) which allows better visualisation of the marker segments. Figure 24 shows what the *Acquisition explorer* should look like after the configuration file has been loaded correctly. (NOTE: This step is not necessary but it does aid in the visualisation of the markers.)

c) Finally, export the file as a TRC file (Select *File* → *Export* → *Motion Analysis Corp.* → *TRC file*).

d) OPTIONAL STEP: *This is an optional step that would be performed before Step 4.5*c*). If the final goal of the entire analysis is to produce a time synchronised joint angle and AE plot then this step should NOT be performed.* Use the frame sliders at the bottom of the window

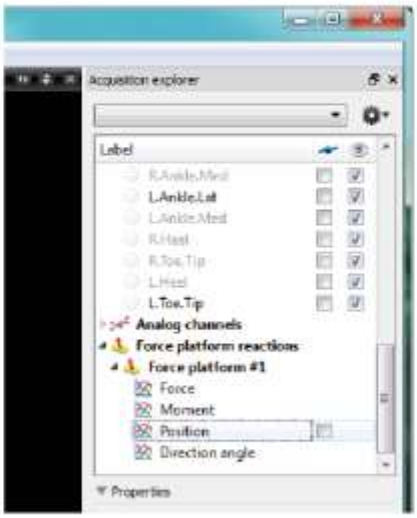

Figure 23: Screenshot of the *Mokka* main window showing the unchecked force position check-box.

to select the *frames of interest*. Then, right click on the frame slider and select *Crop Frames of Interest* (See Figure 25).

5 OpenSim GAIT Analysis

6 Time Synchronisation of AE and IMU Datasets (MATLAB)

7 Time Synchronisation of MoCap and IMU Datasets (MATLAB)

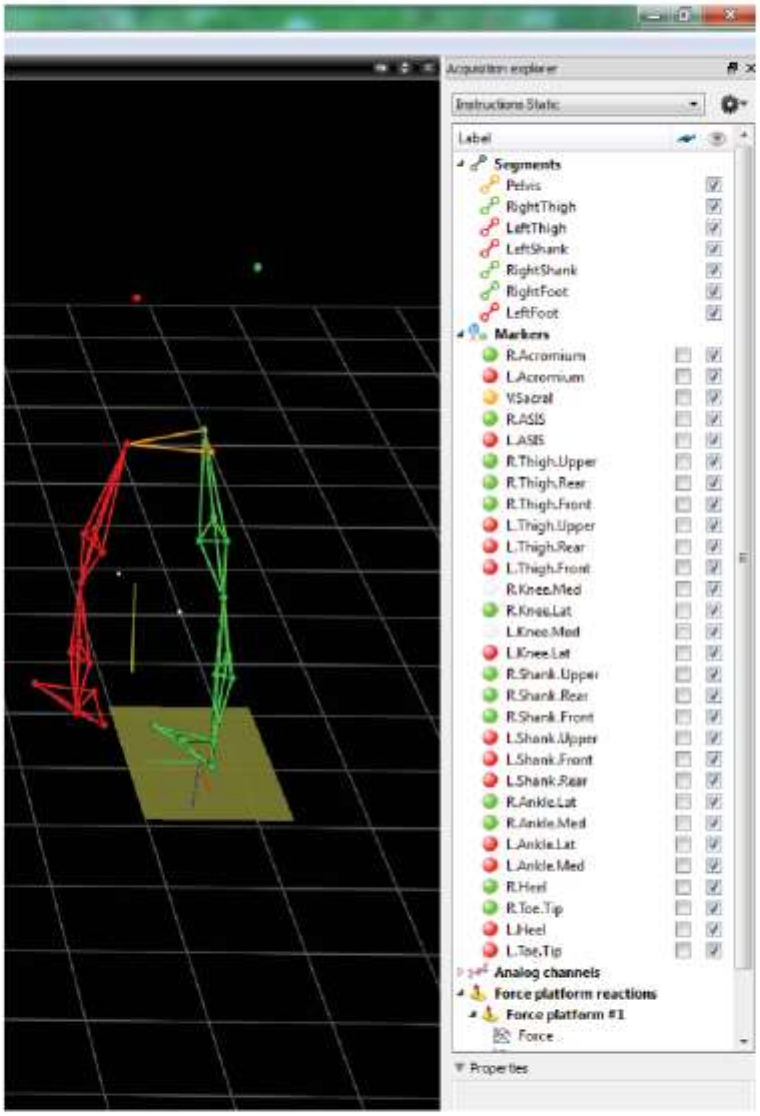

Figure 24: Screenshot of the *Mokka* main window showing the visual configuration after being loaded.

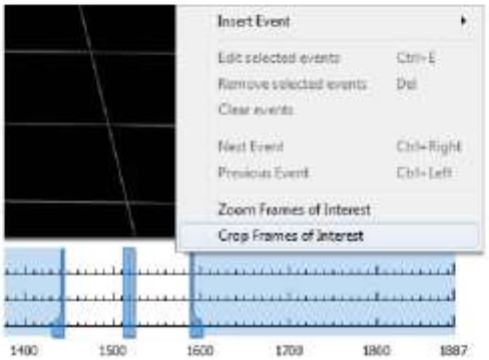

Figure 25: Screenshot of the *Mokka* main window showing cropping to the *frames of interest*.

Appendix D: Motion Capture Instructions

Data Collection Instructions (Motion Capture)

Anthony FitzPatrick

November 2018

1 Introduction

This document outlines the data collection protocol for the 3D motion capture (MoCap) associated with the Hip Assessment and Revision Prediction (HARP) project. The MoCap data collection procedure involves four steps: the recording of anthropometric measurements of the participant, the placement of the reflective markers on the participant, the recording of a static trial, and the recording of dynamic trials. The following sections of the document outline specific information regarding each of the four steps in turn. Note that this document does not describe the set-up and calibration of the MoCap system.

2 Anthropometric Measurements

Some basic anthropometric measurements are taken to assist subsequent scaling of the bones of the OpenSim model which increases the accuracy of 3D motion reconstruction. The basic anthropometrics that need to be measured and recorded are outlined in Table 1. This task is best performed at the beginning of the test procedure before any markers are placed on the participant.

3 Marker Placement

Reflective markers are placed on the participant in predetermined anatomical positions to allow the MoCap system to track body positions. The marker set used is a modified version of that which is included with the OpenSim *Gait 2392* model. The *Gait 2392* marker set has been developed and refined primarily from previous biomechanical studies [1–4] which used 3D motion capture.

Table 1: Required anthropometric measurements

Name	Description
Body Mass	Participant mass.
Height	Participant height.
Inter-ASIS distance	Distance between the left anterior superior iliac spine (ASIS) and right ASIS. Note, this only needs to be measured if markers cannot be placed directly on the ASIS (e.g. in obese participants).
Leg Length (Left & Right)	Full leg length, measured between the ASIS marker and the medial malleolus, via the knee joint. Measure with patient standing, if possible.
Knee width (Left & Right)	The medio-lateral width of the knee across the line of the knee axis. Measure with patient standing, if possible.
Ankle width (Left & Right)	The medio-lateral distance across the malleoli. Measure with patient standing, if possible.

Collectively, the model utilises 39 markers across the lower limbs, pelvis, torso, and head. For the testing described here, only 29 of the model markers will be used with two of these 29 markers being optional.

Table 2 outlines the markers that will be placed on anatomical landmarks. The acromium markers are optional and can be excluded if it is more practicable for a given participant. It is important to place the markers as accurately as possible on the appropriate anatomical landmarks as disparities in positioning adversely effects the results of the OpenSim modelling process. The markers of the knees and ankles are required only for the static trials and can be removed for the dynamic trials.

Table 3 outlines the markers that are placed as part of a marker cluster on either the thigh or shank of the participant. The clusters each contain three markers and hence when attached to a body segment they can be used to track the position of the segment in three dimensions without the need for other markers on the same segment. The exact positioning of the marker clusters is not critical provided that data from a static trial is collected first that also contains knee and ankle markers for reference. The model scaling process of OpenSim allows for the position of the marker clusters, relative to the anatomical markers and hence the bones of the limb they are on, to be determined accurately.

It is important to ensure that the markers are securely fastened to the participant so that they do not move or fall off during the trials. If a marker does move, or it falls off and is repositioned, a

Table 2: Anatomical landmark marker positions

Marker Label	**Definition**	**Marker Position** (* indicates required for static test only)
L.ASIS	Left ASIS	Left anterior superior iliac spine.
R.ASIS	Right ASIS	Right anterior superior iliac spine.
V.Sacral	Sacral	On the skin mid-way between the posterior superior iliac spines (PSIS) and positioned to lie in the plane formed by the ASIS and PSIS points.
L.Knee.Lat	Left Knee Lateral	* On the lateral epicondyle of the left femur and positioned to lie along the estimated axis of rotation of the knee joint.
L.Knee.Med	Left Knee Medial	* On the medial epicondyle of the left femur and positioned to lie along the estimated axis of rotation of the knee joint.
L.Ankle.Lat	Left Ankle Lateral	* On the lateral malleolus of the left shank along the line that would pass through the transmalleolar axis.
L.Ankle.Med	Left Ankle Medial	* On the medial malleolus of the left shank along the line that would pass through the transmalleolar axis.
L.Toe.Tip	Left Toe Tip	If wearing shoes, on top of the shoe at the distal end and halfway between the first and second phalanges. If not wearing shoes, over the distal end of the second metatarsal.
L.Heel	Left Heel	On the posterior-most point of the calcaneus (or the shoe) and at the same height above the plantar surface of the foot as the L.Toe.Tip marker.
R.Knee.Lat	Right Knee Lateral	* Same as L.Knee.Lat but on the right femur.
R.Knee.Med	Right Knee Medial	* Same as L.Knee.Med but on the right femur.
R.Ankle.Lat	Right Ankle Lateral	* Same as L.Ankle.Lat but on the right shank.
R.Ankle.Med	Right Ankle Medial	* Same as L.Ankle.Med but on the right shank.
R.Toe.Tip	Right Toe Tip	Same as L.Toe.Tip but on the right foot.
R.Heel	Right Heel	Same as L.Heel but on the right foot.
L.Acromium (Optional)	Left Acromium	On the left acromio-clavicular joint.
R.Acromium (Optional)	Right Acromium	On the right acromio-clavicular joint.

Table 3: Limb cluster marker positions

Marker Label	Definition	Marker Position
L.Thigh.Upper	Left Thigh Upper	The cluster should be placed so that this marker lies on the lateral thigh and approximately in the coronal plane of the femur. Note that the cluster can be positioned at any distance distally along the femur that is practicable.
L.Thigh.Front	Left Thigh Front	The cluster should be placed so that this marker lies anterior and distal to the L.Thigh.Upper marker.
L.Thigh.Rear	Left Thigh Rear	The cluster should be placed so that this marker lies posterior and distal to the L.Thigh.Upper marker.
L.Shank.Upper	Left Shank Upper	The cluster should be placed so that this marker lies on the lateral shank and approximately in the coronal plane of the tibia. Note that the cluster can be positioned at any distance distally along the shank that is practicable.
L.Shank.Front	Left Shank Front	The cluster should be placed so that this marker lies anterior and distal to the L.Shank.Upper marker.
L.Shank.Rear	Left Shank Rear	The cluster should be placed so that this marker lies posterior and distal to the L.Shank.Upper marker.
R.Thigh.Upper	Right Thigh Upper	Same as the L.Thigh.Upper but on the right thigh.
R.Thigh.Front	Right Thigh Front	Same as the L.Thigh.Front but on the right thigh.
R.Thigh.Rear	Right Thigh Rear	Same as the L.Thigh.Rear but on the right thigh.
R.Shank.Upper	Right Shank Upper	Same as the L.Shank.Upper but on the right shank.
R.Shank.Front	Right Shank Front	Same as the L.Shank.Front but on the right shank.
R.Shank.Rear	Right Shank Rear	Same as the L.Shank.Rear but on the right shank.

new static trial will need to be performed before dynamic trials can be continued. If a new static trial is performed, ensure it is obvious which static trial each dynamic trial is dependent on. It is also important to ensure that the participant's clothing does not move and obscure any markers. It is advisable to tape down any loose clothing that could move to obscure markers. Particular attention should be given to the clothing near the pelvis and thigh markers.

4 Static Trial

A static trial, where marker position data is recorded of the participant standing stationary in an upright position, is carried out to enable the scaling and positioning of virtual markers on the OpenSim model. Once all markers have been placed, the participant stands in the centre of the room (on or near the force plate) so that all markers can be seen by at least two cameras each. It is best if the participant clasps their hands together in front of their chest with the elbows by

their sides in order to ensure the pelvis and thigh markers are not obscured. With the participant in the aforementioned position, approximately two seconds of MoCap data should be recorded and saved with an appropriate file name (e.g. `Static1`). It is advisable to perform multiple static trial recordings (or at least check the recording) to ensure that a recording exists where all 29 markers (or 27 if the optional markers are unused) can be identified in the 3D reconstruction.

It is highly recommended to take photographs from all sides of the participant while they are in the static trial position. The purpose of the photographs is twofold: to confirm the positions of the experimental markers with respect to the participants anatomical landmarks, and to verify the approximate joint angles/positions of the participants limbs during the static trial recording. During the subsequent scaling process of the OpenSim model it is important to check that the positions of the virtual markers on the OpenSim model actually match the positions of the experimental markers on the participant. Additionally, the joint angles/limb positions of the OpenSim model should also match those of the participant during the static trial if the OpenSim scaling process worked well.

5 Dynamic Trials

5.1 Introduction

A number of dynamic trials, where marker position data is recorded concurrently with the acoustic emission (AE) data of the participant performing basic orthopaedic motions, are also carried out to allow joint angle comparisons with AE data. Each dynamic trial begins with a time synchronisation step and then the participant performs a prescribed motion. Prior to the dynamic trials, once a successful static trial has been performed, the knee and ankle markers can be removed from the participant (See the * entries of Table 2) as they are not required for the dynamic trials and removing them simplifies the post-processing of the dynamic MoCap data. However, it is important to note, the accuracy of the subsequent OpenSim modelling will be diminished if any remaining markers are obscured at any point during trials. With the knee and ankle markers removed, more importance is placed on avoiding inadvertent obscuring of any markers from the cameras during dynamic trials and ensuring the cameras lenses were set-up appropriately during calibration prior to all testing.

5.2 Time Synchronisation

All dynamic trials must begin with the time synchronisation procedure. Due to the fact that data is being collected from two different computers (MoCap and AE), a physical time synchronisation event must occur. The AE system contains an additional inertial measurement unit (IMU) which is used to impact the MoCap force plate at the beginning of each dynamic trial. The impact of the IMU onto the force plate produces an acceleration spike in the IMU data which can be matched in time with the corresponding force spike in the force plate data.

A dynamic trial recording should commence with the participant away from the force plate. Once the recording has started, a walking stick with the IMU attached is impacted onto the force plate at least three times with a time interval of at least two seconds between each impact. It is advisable to vary the interval between impacts by a small amount for each separate trial as the post-processing algorithm compares impact interval times to match the corresponding acceleration and force spikes across all dynamic trials.

5.3 Walking

A walking dynamic trial begins with the participant facing the force plate but standing at least two paces away from it. Figure 1 shows a diagram depicting the a good camera layout of the MoCap room and indicates the corresponding walking direction relative to the cameras. Once recording has commenced, the synchronisation procedure of Section 5.2 must be performed. After the synchronisation, the participant walks at a self-selected speed across the room. The participant should walk in such a way that the foot of the leg being acoustically monitored is planted completely on the force plate during one stance phase of the walk across the room. The participant should also not attempt to stop walking until the foot they planted on the force plate has completed the subsequent swing phase. The MoCap recording can be stopped once the participant has completed the walk across the room.

5.4 Other Motions

As per the AE data collection protocol, the participant is asked to perform other motions. The other motions that the participant is asked to perform are squats, standing up from and sitting down

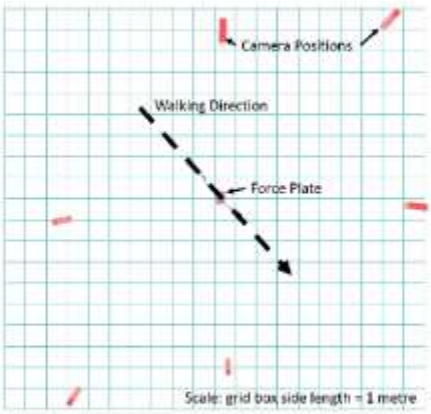

Figure 1: Diagram depicting MoCap camera layout and walking direction.

onto a chair, step-ladder ascents and descents (if a step-ladder is present), heel-strikes (impacting the force plate with the heel of the leg being acoustically monitored), and any other motions that induce implant noises. Each of these recordings must also commence with the synchronisation procedure of Section 5.2.

6 References

[1] Samuel R. Hamner, Ajay Seth, and Scott L. Delp. "Muscle contributions to propulsion and support during running". In: *Journal of Biomechanics* 43.14 (2010), pp. 2709–2716.

[2] A Cappozzo et al. "Position and orientation in space of bones during movement: anatomical frame definition and determination". In: *Clinical Biomechanics* 10.4 (1995), pp. 171–178.

[3] M P Kadaba, H K Ramakrishnan, and M E Wootten. "Measurement of lower extremity kinematics during level walking". In: *Journal of Orthopaedic Research* 8.3 (Nov. 1990), pp. 383–392.

[4] Roy B. Davis et al. "A gait analysis data collection and reduction technique". In: *Human Movement Science* 10.5 (Oct. 1991), pp. 575–587.